

AN AUTOSEGMENTAL/METRICAL ANALYSIS OF SERBO-CROATIAN INTONATION *

Svetlana Godjevac

Abstract

Based on the qualitative analysis of the F_0 contours of wide range of utterances (broad focus declaratives, broad focus questions, narrow focus declaratives, narrow focus questions, vocative chant, and prompting intonation) uttered by nine native speakers, an autosegmental/metrical analysis of Standard Serbo-Croatian intonation is proposed. This analysis argues for sparse specification of tones, contra Inkelas and Zec (1988), and two levels of prosodic phrasing: the phonological word and the intonational phrase. The phonological word is defined in terms of a lexical pitch accent and an initial word boundary tone, whereas the intonational phrase is a domain defined by pitch range manipulations (expansion, compression, reset, downstep) and final intonational phrase boundary tones.

1 Introduction

Standard Serbo-Croatian (SC) is a pitch-accent language. All analyses (Browne & McCawley 1965 (B&M), Inkelas & Zec 1988 (I&Z), Kostić 1983, Lehiste & Ivić 1963,

*I would like to express my gratitude to Mary Beckman, Chris Barker, Allison Blodgett, Rebecca Herman, Molly Homer, Tsan Huang, Ilse Lehiste, Gina Taranto, and Pauline Welby. I also wish to thank my informants: Dragana Aleksić, Ljubomir Bjelica, Ana Dević, Ksenija Djuranović, Svetislav Jovanović, Jasna Kragalott, Svetlana Likić, and Branislav Unković for their patience and kindness in providing the data. All errors are mine.

1986 (L&I), Nikolić 1970, Stevanović 1989, Gvozdanović (1980), *inter alia*) recognize four different types of accents: short falling, long falling, short rising, and long rising. In this paper I present an analysis of surface tones of these accent types in different sentential environments, including broad-focus and narrow-focus utterances, citation form, vocative chant, prompting intonation, and questions. This analysis is based on the instrumental study of recorded utterances by eight native speakers. It is an autosegmental/metrical analysis because the F_0 shapes are decomposed into their component parts, and the tones and the backdrop pitch range are analyzed in terms of their relations to metrical structure.

The general observation differentiating this proposal from earlier autosegmental accounts, is that even the surface tones in SC are sparsely specified to moras, the tone bearing unit in SC. More specifically, the analysis argues for three main innovations over the cited analyses: (i) a decomposition of word tone strings into a demarcative tone, a boundary tone, and accent proper (rather than H-tone spreading and default L-insertion); (ii) bitonal accents with the initial tone starred (i.e. associated to the accented syllable) and the trailing tone unassociated; and (iii) no neutralization of the lexical accents in declarative sentence final position. My proposal regarding SC prosodic structure includes two prosodic units: a phonological word and an intonational phrase. Their tonal properties are defined in terms of specification of accents, boundary tones, and pitch-range manipulation. In addition, some observations of more global pitch trends, such as downstep, are offered.

One reason a refined picture of SC word tones is important is that it serves as the foundation of an ongoing study of the interaction of intonational effects such as pitch range compression and downstep with syntactic scrambling, word-order focus, etc. These in turn are central to interpretation. The interaction of intonation with interpretation is left for future study. More immediately, this study serves to add to descriptions of prosodic structure of pitch accent languages, which include Japanese, Norwegian, and Swedish, thereby contributing to the crosslinguistic study of variation in prosody.

I argue that SC's four accents are bitonal. The falling accents are H^*+L , whereas the rising accents are L^*+H , where '*' marks the tone associated with the relevant tone bearing unit within the stressed syllable, as in Bruce's (1977, 1990) analysis of Swedish word accent. The consequence of this proposal is that the second tone is not linked to a particular mora but is phonologically unassociated. As we will see, a long falling accent may realize both tones on the stressed syllable, whereas in words with a short falling accent, the trailing tone is usually realized on the poststressed syllable, and sometimes is even truncated.

Not all words in SC carry a pitch-accent. Verbal and pronominal clitics, prepositions, and most conjunctions do not bear pitch-accents.¹ These words cliticize to an adjacent word which does bear a pitch-accent to form a phonological word. A phonological

¹Zec & Inkelas (1990) assume that the division between phonological words and clitics aligns with the syntactic division into content and function words. This division seems generally right but there are a few

word is the smallest prosodic unit, and is tonally marked by a pitch accent. (As we'll see, proclitics are marked by a L word boundary tone (which I will mark as %L), but they lack a pitch accent.) In the case of SC, I propose, the relevant tonal marking is a pitch-accent and a %L word boundary tone. As a general rule, there is maximally one pitch accent and one %L word boundary tone per phonological word. (As will be discussed in section 4.1.2, there are exceptions to this rule. Some polymorphemic words can be realized with two pitch-accents, but they are in free variation with variants realized with one pitch accent. Proclitics also bring an additional word boundary tone.) That is, a phonological word in SC has exactly one head syllable (marked with lexical pitch-accent) and at least one edge (word boundary) marked tonally.

The sentential tune in a declarative utterance under broad focus shows an overall downtrend in the pitch level. (By broad focus I mean the sentential tune which lacks prosodic focus. Prosodic focus will be discussed in section 4.2.4.) My as yet unquantified observations of many F_0 contours suggest that much of this downtrend can be described as a downstep at each word boundary. That is, the word boundary tone downsteps the succeeding H target. The final constituent in a sentence then usually ends up in a lower pitch range than any other constituent in the sentence. This cues the end of the sentence. On the basis of instrumental evidence, L&I have pointed out the potential for neutralization of word accents in disyllabic words in this position. I&Z have characterized this phenomenon by the phonological rule of L insertion whose effect is to erase the tonal lexical distinctions. However, I show, using minimal pairs, that the lexical tones are still present in this position despite the smaller range for their manifestation (see, sections 4.2.2 and 4.2.4). Therefore, I conclude that a different phonological model is needed from the one I&Z propose. The new model needs to be able to separate the effects of the gradient backdrop pitch trends from categorical tone deletion.

As for sentence-level prosody, words under prosodic focus, in narrow focus utterances, show a higher target for the accent H relative to the same utterance without the prosodic focus. This is true both for the starred tone of the falling accents (H^*+L) and the trailing tone of the rising accents (L^*+H).

In summary, in this paper I posit three prosodic units for Serbo-Croatian: a phonological word, an intermediate phrase, and an intonational phrase. The declarative sentence pattern of SC shows a continuous alternation between H and L tones. Every phonological word is marked by this pattern, and so is each sentential string. However, the sentence intonation is more than just a concatenation of the word accent tones. The declarative sentence intonation can be accounted for by positing a word-boundary tone, a downstep rule phrase internally, the rule of reduction of pitch range in final position, super H targets for

exceptions. For example, demonstrative pronouns, which function as determiners, thus function words, do bear a pitch-accent. Some conjunctions, such as *pā* 'so', *iako* 'although', *ali* 'but', etc. also bear an accent.

discourse-initial segments, and pitch range reduction of post-focus positions. On the basis of contrasts in different melodies, such as declaratives, prompting intonation, and vocative chant, I argue for two different tonal markings of an intonational phrase: two boundary tones, L%, H%, and two phrase accents L- and H- for an intermediate phrase.

The paper consists of three major parts. The first major part, section 3, deals with lexical accents and their properties; the second part, section 4, is concerned with tonal markings of prosodic structure; and section 5 deals with the issues of interaction between the lexical and structural markings. Section 6 concludes by summarizing the proposed analysis of Serbo-Croatian intonation.

2 Methodology

The language that I intend to cover in this paper is the Štokavian–Ekavski variant² Standard SC. The analysis presented here is a broad outline investigation. It is based on an instrumental investigation of F₀ contours for close to 300 utterance types, ranging from citation form utterances of single words to three-sentence paragraphs. The intention was to provide a wide coverage of Serbo-Croatian utterance types in order to get an overview of the complete system, as a framework for investigating some specific aspect of the system in a thorough quantitative analysis with careful control of interaction with other sources of systematic variation. This purpose is a result of the need for the more overall picture of the system prior to the later quantitative modelling of specific questions. This is in line with the work done by Pierrehumbert (1980), which provided the groundwork of a complete description of the English intonational system, and which subsequently resulted in the detailed study of pitch range in Liberman & Pierrehumbert (1984). Consequently, results presented here will be more suggestive than quantitative.

All the material uttered by the author was digitally recorded directly into a Sun workstation (Sun4) or Linux box and analyzed using the Entropics Waves program. Materials uttered by the other seven native speakers were recorded in a quiet room on a Marantz taperecorder and then digitized with Waves using a Denon tape player and the Sun workstation. Four of the speakers, including the author, are from Novi Sad, three of the speakers are from Belgrade, and one of them is from Kruševac.

For the purposes of getting an uninterrupted pitch track, almost all of the words and sentences recorded were chosen for their all-sonorant quality. Some exceptions were made when the length or the late position of the accent of the word was crucial in investigating a certain hypothesis and no word with all sonorants was found with those characteristics.

²Serbo-Croatian dialects are divided along two parameters: (a) the first parameter is the word for 'what', thus we have *što*, *ča*, and *kaj* and the corresponding dialects: Štokavian, Čakavian, and Kajkavian; (b) the second parameter is the reflex of the Old Church Slavic vowel *jat*. There are three reflexes of this vowel: [e], [i], and [ije]. Hence the corresponding dialects: Ekavski, Ikavski, and Ijekavski.

SERBO-CROATIAN INTONATION

Also, as it was important to look at minimal pairs and words with particular syntactic and semantic properties (notably, *wh*-words) it was necessary to include some words that do not have all-sonorant quality.

All the pitch tracks in this paper are utterances performed by the author. This decision is a consequence of the fact that it was not possible to get all the relevant data from all the speakers, and was used to keep the pitch contours consistent throughout the paper for ease of comparison. However, none of the pitch tracks used here for the purpose of illustration are isolated tokens of the type. Pitch tracks were used as evidence only when the same contour occurred constantly across at least five tokens of the same type of utterance.

3 Lexical Information

3.1 Lexical Accents and Their Distribution

The standard description of the distribution of the accents is that falling accents only occur on the initial syllable and that rising accents occur on any syllable but the last syllable. Thus, rising accents never occur in monosyllabic words since the initial syllable is also the last syllable. So, monosyllabic words necessarily have a falling accent. The traditional way of marking falling accents is: [\sim] for the short falling, and [\wedge] for the long falling. Some examples of words with these accents and a pitch track of a word under the accent in a sentence medial position are given in the following table.

SVETLANA GODJEVAC

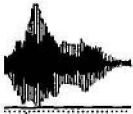

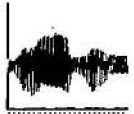
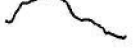
SHORT-FALLING	LONG-FALLING
  jǎlov 'fruitless'	  jâvan 'public'
lāv 'lion' jǎlov 'fruitless' lūla 'pipe' nǎjava 'announcement' nēminovan 'inevitable' nēravnomeran 'uneven' paradāajz 'tomato' ōrṇamentalan ranorānilac	lāž 'a lie' jâvan 'public' ūlje 'oil' nāmēra 'intention' vōljan 'willing' vāljan 'rolled' ūman 'wise' ūmoran 'tired' rēvija 'review'

Table 1: The F₀ tracks show the two falling accents in words *jǎlov* 'fruitless' and *jâvan* 'public' in a sentence medial position to circumvent discourse or sentence edge effects. The rest of the table provides examples of words under the two falling accents, short and long, with the stress on the first syllable, differing in length.

The traditional way of marking the rising accents is the following: [`] for the short rising, and [´] for the long rising. A rising accent can occur on any syllable but the last and it never occurs on monosyllabic words. Here are some examples:

SERBO-CROATIAN INTONATION





SHORT-RISING	LONG-RISING
  <p>màlina 'raspberry'</p>	  <p>línija 'line'</p>
nàum 'intention' màlina 'raspberry' ìronija 'irony' mànjiina 'minority' ànemija 'anemia' artiljèrija 'artillery' neprikosnòven 'sacred' multimiliòner 'multimillionaire'	jáje 'egg' línija 'line' živeo 'he lived' aróma 'aroma' nenámeran 'unintentional' memoári 'memoirs' anulírala 'she annulled' legitimácija 'ID' nacionalizácija 'nationalization'

Table 2: The F_0 tracks show the two rising accents, in words *màlina* 'raspberry' and *línija* 'line', in a sentence medial position to circumvent discourse or sentence edge effects. The rest of the table provides examples of words under the two rising accents, short and long, with varying position of the stress, and differing in length.

3.2 Phonemic Distinctions

Both distinctions, short vs. long and falling vs. rising contrast words. The following sets of minimal pairs show the contrastive role these properties play:

SHORT-FALLING	LONG-FALLING
râd 'eager'	râd 'work'
sâd 'now'	sâd 'plantation'
ôran 'plowed' (participle)	ôran 'disposed'

Table 3: Examples of minimal pairs for the two falling accents.

SVETLANA GODJEVAC

SHORT-FALLING	SHORT-RISING
mòli '(s)he begs'	mòli 'Beg!' (imperative)
òpasan 'dangerous'	òpasan 'with a belt'
òrao 'he plowed'	òrao 'eagle'
òsnovan 'founded'	òsnovan 'elementary'

Table 4: Examples minimal pairs for the two short accents.

LONG-FALLING	LONG-RISING
rāvan 'plain'	rāvan 'flat'
rādi '(s)he works'	Rādi 'to Rada'
nēma 'he doesn't have'	nēma 'deaf.fem'

Table 5: Minimal pairs for the two long accents.

SHORT-RISING	LONG-RISING
sèdeti 'to sit'	sédeti 'to go gray'
òpisan 'described'	ópisan 'descriptive'
rásipan 'wasted'	rásipan 'wasteful'

Table 6: Minimal pairs for the two rising accents.

There are also minimal pairs that cut across both dimensions. That is, words with the same segmental tier but with tonal contrast along both long/short and rising/falling parameter:

LONG-RISING	SHORT-FALLING
(h)rāna 'food'	rāna 'wound'
LONG-FALLING	SHORT-RISING
vâljan 'rolled'	vâljan 'good'

Table 7: Minimal pairs for both duration and pitch oppositions.

3.3 Lexical Tones

Serbo-Croatian pitch-accent can be characterized by the position of stress and the specification of two tone levels, high and low, as already proposed by I&Z in the framework of autosegmental phonology and earlier by Halle (1971). The order and distribution of these

SERBO-CROATIAN INTONATION

tones relative to the accented syllable corresponds with the type of the lexical accent for which the word is specified.

It has been claimed that the distinction between the rising and the falling accents lies in the fact that the rising accents are bisyllabic whereas falling accents are monosyllabic:

'... relying on perceptual evidence analyzes rising accents as encompassing two syllables ... However, falling accents encompass only one syllable.' (B&M:147 citing Hodge 1958, Bidwell 1963, Masing 1876, Ivić 1958, 1961).

'All four accents have traditionally been treated as associated with a single syllable, as the diacritics [...] show. However, only the falling accents are clearly monosyllabic; the rising accents are disyllabic in nature, as we will see.' (I&Z 1988:227, footnote 2.)

This distinction suggests the assumption that since accent is (by definition) a culminative marker within its domain, the relevant phonetic property should culminate at the accent location; hence a 'pitch accent' should be a pitch culmination, i.e. a peak in the pitch contour localized at the accent. The falling accents are in accordance with this assumption since the characteristic of the falling accents is that the H tone is realized on the accented syllable itself. The rising accents, on the other hand, deviate from this common assumption about alignment between accents and peaks. The H of the rising accents is realized on the post-stressed syllable. This misalignment between the accented syllable and the peak in the rising accents has thus far been couched in terms of durational properties of the accent: monosyllabic, vs. bisyllabic.

Instead of thinking of the two classes of SC pitch-accent, falling vs. rising, in terms of monosyllabic vs. bisyllabic accents, I propose to switch the perspective from the number of syllables necessary to realize the accent peak to thinking of the number of tone targets necessary to realize a rise or a fall, i.e. to consider both of them as being bitonal, where only one of the tones is anchored to a stressed syllable (cf. Bruce 1990). For the falling accents, the anchored tone will be the H, and for the rising accents the anchored tone will be the L. The data show that the second (trailing) tone can be realized on the stressed syllable as well, as in the case of the long-falling accent, but it is usually on the poststressed syllable, as is the case for all other accent types. Consequently, it seems more appropriate to treat the second tone as unassociated rather than anchored to a particular syllable or mora.

This view is more in accordance with the position in Kostić (1983) who argues that all four accents should be treated as bisyllabic. His argument involves the claim that accent peaks are fully realized only in opposition to the following or preceding syllable. However, in order to claim this, he has to exclude monosyllables, which he then treats as exceptions. The position taken in this paper is that pitch accents are only partially linked to a particular

metrical position. Only the first tone of the (bitonal) accent is anchored to the text, and this is the starred tone.

The next three sections explore the consequences of this proposal for the falling accents alone, the rising accents alone, and both together.

3.3.1 Falling Accents

In a declarative utterance, falling accents can be characterized in terms of two tones, H followed by L. Both short-falling and long-falling accents have a H tone on the stressed syllable followed by a L tone. The difference between the two accents seems to be not only in the duration of the syllable under stress but also in the timing of the tonal qualities. In words with long-falling accents the L seems to show up during the stressed syllable whereas in words with the short-falling accent the L starts after the stressed syllable.

Here are some examples of citation forms. Figures 1 through 3 show similar (or minimally contrasting) words, with one, two or three syllables. In all of the figures throughout the paper, the cursors (vertical lines) mark the ends of syllables (or the end of words, when individual syllables are not marked).

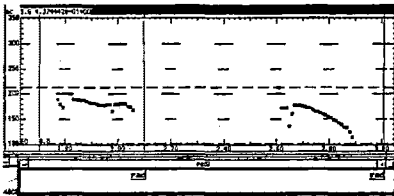


Figure 1 : Short-falling vs. Long-falling, 1 syllable words. The first utterance, on the left, is the word: *rād* 'eager', the short-falling accent; the second utterance is the word *rād* 'work', the long-falling accent.

SERBO-CROATIAN INTONATION

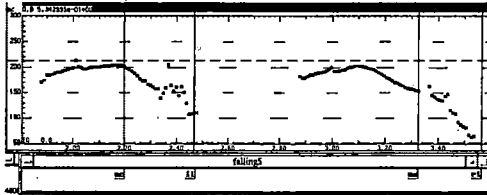


Figure 2 : Short-falling vs. Long-falling, two-syllable words, (citation form). On the left is the word *mōli* 'begs', short-falling accent; on the right is the word *māri* 'cares', long-falling accent.

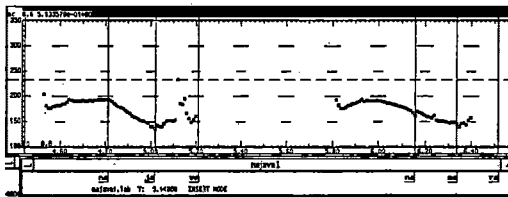


Figure 3 : Short-falling vs. Long-falling, 3 syllable words. The first utterance, on the left, is the word: *nājava* 'announcement', the short-falling accent. ; the second utterance is the word *nāmera* 'intention', the long-falling accent.

From the three pitch tracks above we can notice that the two accents are associated with a particular shape of F_0 . In the long-falling examples, the accented syllable carries both H and L tone, whereas in the short-falling examples, the accented syllable carries only the H tone, and the L tone is realized on the post-stressed syllable. In other words, the alignment of the peak in the long-falling accent is more towards the middle of the syllable, whereas in the short falling-accent it is at the right edge of the syllable.

There is a difference between monosyllables (Figure 1) and disyllabic words (Figure 2), under the short-falling accent. The disyllabic words show the L tone on the post-stressed syllable whereas in the monosyllables the fall is truncated when the word is in isolation.

Despite this difference in monosyllables, the two falling accents are very similar. Consequently, I propose that these accents be represented as H^*+L . This representation accurately captures the fact that the H tone is anchored (associated) to the stressed syllable,

whereas the L tone is a trailing tone, which may or may not fall on the stressed syllable, or may even be truncated. The distinguishing property between the two is the duration of the anchored tone. The duration of the H tone is shorter in the long-falling accent than in the short-falling accent. That is, the fall from H to L starts earlier in the long-falling accent than in the short-falling accent. Hence the steepness of the fall differs between the long-falling and the short-falling accents. This produces the effect of the long-fall (shorter H tone) vs. short-fall (longer H tone). Hence the name that they bear seems clearly appropriate. The duration of the vowel under the two falling accents is not always the best cue for which type of accent we are dealing with. The durations of the H tones seem to be more distinct than the durations of the vowels. Although there is a contrast between short and long vowels in unstressed positions, duration is the best cue for stress in SC, as shown by L&I. That is, a short stressed vowel is longer than a short unstressed vowel, and a long stressed vowel is longer than a long unstressed vowel.

3.3.2 Rising Accents

In a declarative utterance a rising accent exhibits a L tone on the stressed syllable and a H tone on the post-stressed syllable. There does not seem to be an obvious qualitative difference in F_0 between the two rising accents analogous to the steepness of the fall or the length of the starred tone in the falling accents. Also the difference in F_0 target is insignificant, when we compare either the peaks or the preceding lows. This is also confirmed by the data reported in L&I (1985). However, there is a difference in vowel quality. The vowels under the long-rise are more peripheral than the vowels under the short-rise. This may be in part due to the difference in duration of the vowel, since the vowel in a word with the long-rising accent is longer than in a word with the short-rising accent. As L&I (1963:93f) report, the long /e/, /o/, and /a/ are more peripheral than their short allophones.

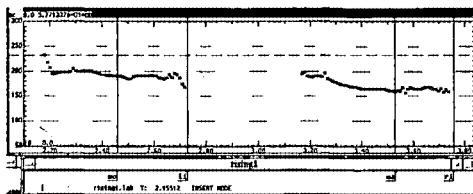


Figure 4: Short-rising vs. Long-rising, two-syllable words (citation forms). The first utterance is the word *mòli* 'beg!' (imperative), short-rising accent; the second utterance is the word *Mári* 'to Mara', long-rising accent.

SERBO-CROATIAN INTONATION

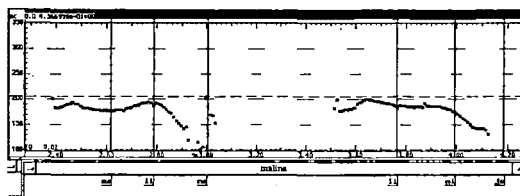


Figure 5: Short-rising vs. long-rising, 3 syllable words. The first utterance is the word *málina* 'raspberry', the short-rising accent; the second utterance is the word *línija* 'line', the long-rising accent.

In the above figures we can see that the H of the rising accents is not very prominent. This is because these short utterances are citation forms, which inevitably encompass the phenomenon of final lowering (to be discussed in section 4.2.3). For the purpose of the illustration of this effect I present the word *omalovažavanje* 'humiliation' in two different environments, a citation form and as an initial constituent of a sentence.

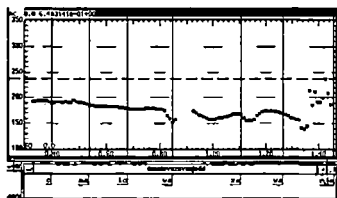


Figure 6: Short-rising accent. The utterance is a citation form of the word *omalovažavanje* 'humiliation'.

This word was chosen for its late accent placement, which allows a long stretch of syllables before the accent. We can notice the L tone, which is anchored to the stressed syllable *-žà-*. The post-stressed syllable is the one that the H tone usually gets realized on. Consequently, the choice for the representation of this type of accent is L*+H. In this case, that is, the citation form, the H tone is affected by the discourse final position, i.e. final lowering. This is the effect that L&I called neutralization of the accents in the final position. However, as will be discussed in section 4.2.2 this effect is due to the pitch-range reduction, and comparison to falling accents clearly shows the preservation of the distinction between the two types of accents. Since utterance final elements are affected by the position, it is especially illuminating to compare the citation form with a non-citation

form. To anticipate the discussion of accents in contexts of sentences, in section 4.2, I show the same word in a sentence initial position in the following figure.

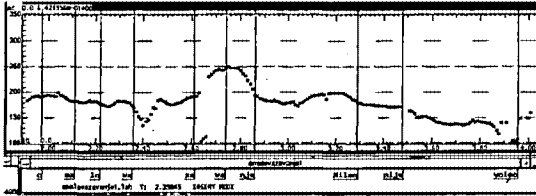


Figure 7: **Short-rising accent.** This is a sentence-initial utterance of the word *omaložavanje* 'humiliation' ('Humiliation, Milan didn't like.').

As we can see in Figure 7, the H tone of the rising accent is more visible due to the continuation of the utterance. I will return to the sentence level influence on the accents in sections 4 and 5. In the next section I continue to discuss the properties of the lexical tones. The point of interest here are distinctions among accents.

3.3.3 Rising/Falling Opposition

In the previous section the opposition between the short and long accents was shown. The F_0 contour very clearly reflects the opposition between the falling accents. For the rising accents, the F_0 is a less transparent indicator of the contrast between the short and the long rising accent. It is the time course of the H^* and the steepness of the fall that create a distinction between the falling accents. The rising accents, on the other hand, do not seem to have as clear a tonal distinction, in terms of the F_0 manifestation: rather they differ in vowel quality.

In this section I present the opposition between the rising and falling accents of the same durational type because this allows us to see the difference between a rise and a fall most clearly, since the length variable is kept constant. Figure 8 shows a minimal pair, the long falling vs. the long rising accent; Figure 9 shows a minimal pair for the short falling/rising opposition.

SERBO-CROATIAN INTONATION

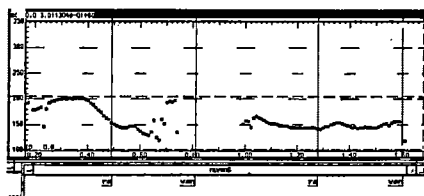


Figure 8: Long accents. The first utterance is the word *râvan* 'plain', the long-falling accent; the second utterance is the word *râvan* 'flat', the long-rising accent.

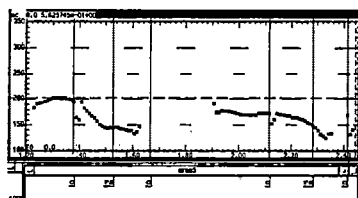


Figure 9: Short accents. The first utterance is the word *ðrao* 'he plowed', the short-falling accent; the second utterance is the word *ðrao* 'eagle', the short-rising accent.

The difference between the falling and the rising accents is very clear from the above pitch tracks. The falling accents exhibit a clear fall in the pitch, whereas the rising accents exhibit a small rise or a steady pitch on the post-stressed syllable. The lack of an obvious rise, i.e. a clear manifestation of the H target, in these examples is due to the citation form intonation of the utterances. As we saw in the preceding section, the rising accents do realize the H tone, which is higher from the tone of the stressed syllable, as long as the word is not utterance final. In addition, we can see that the H tone of the falling accents is considerably higher from the H tone of the rising accents. (This is not an artifact of their order in the list since, the reversal of their linear order in production produces the same effect; see Figures 8, and 9.)

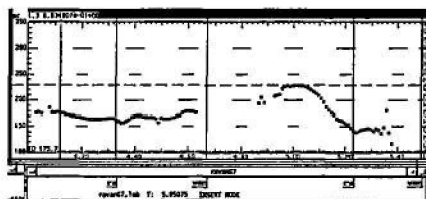


Figure 10: Long accents. The first utterance is the word *ravan* 'flat', long-rising accent; the second utterance is the word, the *ravan* 'plain', the long-falling accent.

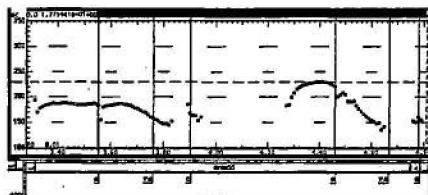


Figure 11: Short accents. The first utterance is the word *đrao* 'eagle', the short-rising accent; the second utterance is the word *đrao* 'he plowed', the short-falling accent.

This observation has been noted by Kostić (1983) and the F_0 measurements for 3 different pitch ranges of speakers, low, medium and high, from L&I (1963) also support that conclusion.

3.3.4 Lexical Tone Analysis

In this section I give a proposal for analyzing the four lexical tones of the four Serbo-Croatian pitch accents.

As previously mentioned, the falling accents are characterized by the HL melody, whereas the rising accents exhibit LH melody. The tonal distinction on the short-long parameter is manifested with the falling accents (in the steepness of the fall), but not with the rising accents. Schematically, however, the proposal for the four accents can be represented as the following:

SERBO-CROATIAN INTONATION

	FALLING	RISING
SHORT	$\sigma \quad \sigma \quad \sigma$ $ \quad \quad $ $\mu \quad \mu \quad \mu$ $ $ H^*+L	$\sigma \quad \sigma \quad \sigma$ $ \quad \quad $ $\mu \quad \mu \quad \mu$ $ $ L^*+H
LONG	$\sigma \quad \sigma \quad \sigma$ $/ \backslash \quad \quad $ $\mu \quad \mu \quad \mu \quad \mu$ $ $ H^*+L	$\sigma \quad \sigma \quad \sigma$ $/ \backslash \quad \quad $ $\mu \quad \mu \quad \mu \quad \mu$ $ $ L^*+H

Table 8: Surface representation of tones in trisyllabic words with the stress on the first syllable.

In the above graphs the distinctions between falling and rising accents is represented by the HL and LH melodies, whereas the short/long distinction is captured by the mono-moraic vs. bi-moraic status of the syllable to which the accent is associated. So, all the accents are bitonal: however, of the long accent types, only the first tone is anchored to the first mora of the falling accents and the second mora of the rising accents. (The justification for the particular anchoring site within the syllable for the long rising accent requires explanation of one of the phrasal tones and is deferred until section 4.1.)

This differs from the analysis proposed by Inkelas & Zec (1988) (I&Z) in two important ways. First, in their autosegmental analysis, all tone bearing units are specified for tone at the surface. Hence in their theory the structural fact of accent is only a property of an underlying form, whereas in this analysis the accent is viewed as a pitch event localized at the stressed syllable. Second, in I&Z's theory the rising accents are represented as a sequence of two H tones, whereas in this analysis, the rising accents are a LH melody, where the L tone is anchored to the stressed syllable.

The two analysis agree on the representation of the difference between long and short syllables through a moraic structure. For the sake of comparison, I provide a schema of their analysis of trisyllabic words with an accent on the initial syllable after the derivation is completed:

	FALLING	RISING
SHORT	σ σ σ μ μ μ H L L	σ σ σ μ μ μ H H L
LONG	σ σ σ /\ μ μ μ μ H L L L	σ σ σ /\ μ μ μ μ L H H L

Table 9: Predictions of the surface representation of tones in trisyllabic words with the stress on the first syllable, according to Inkelas and Zec (1988). (Compare to Table 8)

Under their analysis, the HL melody of the long-falling accent is realized on the two moras of the accented syllable itself, whereas the short-falling accent is realized across two syllables. In the case of the rising accent, which they represent as HH, the two H tones are associated to the last mora of the accented syllable and the first mora of the post-stressed syllable. This is because in their theory, two adjacent H tones cannot belong to the same syllable. According to my data, both rising accents have the H tone realized on the post-stressed syllable only. That is, the high tone is never realized on the accented syllable or the last mora of the accented syllable. To make the point clearer, I will present the instrumental data of the examples analyzed in their paper, and discuss the predictions their analysis makes about the surface tones.

The following figure shows a pitch track of the two rising accents discussed and analyzed in I&Z:

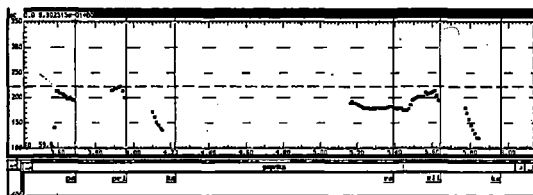


Figure 12: Figure of two rising accents: *pàprika* 'pepper' and *rázlika* 'difference'.

SERBO-CROATIAN INTONATION

For the purpose of comparing I&Z's analysis to the one presented here, we must abstract away from the basic difference between these two analyses, such as full vs. sparse surface specification of tones. With that in mind, we can compare only the accent representation in the two approaches.

According to their analysis, the difference between *pàprika* and *rázlika* is in the position of the two H tones: *rázlika* should realize the H tone on the second mora of the first syllable, (*rá-*) and the first mora of the second syllable (*-zli-*), whereas *pàprika* should have a H tone on the first syllable (*pa-*) and the second syllable (*-pri-*). However, we can see that the H tone peak is always realized on the post-stressed syllable. Also, the accented syllable has a L tone in both of these accents. The following pitch tracks also confirm this observation. Figures 13 and 14 show words with long rising and short rising accent, respectively, on the third syllable in a five-syllable word. Both words are uttered in a sentence medial position of a broad focus utterance.

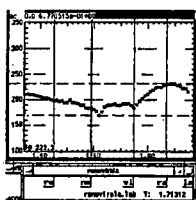


Figure 13: The word *renovirala* 'she renovated' in utterance medial position.

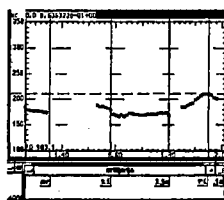


Figure 14: The word *artiljerija* 'artillery' in utterance medial position.

As Figures 13 and 14 show, the H tone is a property of the post-stressed syllable and it is not shared by the two consecutive moras of the stressed and post-stressed syllable. Thus, I&Z's hypothesis is not consistent with the instrumental data. On the basis of the instrumental evidence both from my corpus and from the corpus presented in L&I, I assume that the H tone of the rising accents is a property of only the post-stressed syllable.

Another clear advantage of assuming that only one tone of these accents is anchored to the accented syllable involves the treatment of monosyllabic words. As we saw earlier, in section 3.1, falling accents can occur on monosyllabic words. If we assume that tones are properties of moras and are anchored in them, then only the long-falling accent would be able to occur on monosyllabic words, since long syllables have two moras and the two tones, H and L would be associated with them. However, the short-falling accent, which is a property of short syllables, thus only one mora is available, would have no place for the L tone. Under my analysis, the fact that both types of falling accents are present in the language falls out as a natural consequence of the fact that only the H tone is anchored to

the stressed syllable. The L tone is capable of being truncated if there is only one syllable for realization of tones.

So, we see that having a representation which anchors only one of the tones to a particular syllable gives us a natural explanation for why monosyllables can have both short and long falling accent. In the theory of I&Z, this fact is not accounted for. However, this same reasoning should give us an explanation for the other part of the distributional fact of the SC accents (that is, an explanation for why rising accents never occur on monosyllables). The traditional explanation has always resorted to the idea that the rising accents are bisyllabic, unlike the falling accents. This is a restatement that still calls for an explanation. But, as I have tried to show in this section: In non-monosyllables, the short-falling accent can also be characterized as bisyllabic. So, is there a natural explanation for the distributional properties of the rising accents within this system? I believe there is. The explanation, offered in the next section, involves reasoning about the functional properties of tones and how densely-distributed similar tones can realize their functions. But, before we can go to that explanation (see section 4.1), it is first necessary to introduce another property of the SC prosodic word, a L word-boundary. The presence of the word boundary tone is more prominent in utterances that consist of more than a single word, thus we turn to the sentence-level tonal properties of the Serbo-Croatian prosody.

4 Structural Information

4.1 Phonological Word

In this section I define the smallest prosodic unit in SC, the phonological word. I show that tonal markings of this prosodic unit are of two types: one demarcative (a left edge tone) and the other culminative (the pitch accent).

4.1.1 Word-Boundary Tone

In addition to the lexical tones considered to be realizations of the word accent type, each phonological word in Serbo-Croatian exhibits a boundary tone as well. That is, each word that bears an accent must have a L boundary tone, which I represent as %L. I will argue that this tone always precedes the lexical tonal realizations for reasons that will be clearer when the discussion of downtrend gets introduced.

In their autosegmental account of Serbo-Croatian tones, I&Z assume that words are specified for the H tones in the lexicon, whereas L tones are assigned late in the process of derivation. For declarative intonation they propose a classical tone association account whereby at the end of a derivation each mora is associated to exactly one tone, either the accent H or the default L. So for example, in a disyllabic word with a long-rising accent

only the first mora of the first syllable would be assigned the L tone, since the H tone would be assigned to the second mora of the first syllable and the single mora of the second syllable (if it is not phonemically long) in accordance with their derivational rules. A disyllabic word under a long-falling accent would have the first mora of the first syllable associated to a H tone. A string which includes these two types of words would be a good test for word boundary tones, since the I&Z theory predicts no other tones besides the lexical tones. This string would be predicted to have the following tone contour:

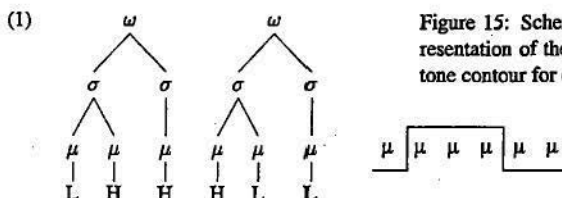


Figure 15: Schematic representation of the expected tone contour for (1).

For the purpose of testing this prediction, I have constructed three types of examples: (a) a sequence of a noun subject and a verb (since SC is an SVO language, this sequence does not involve any type of pragmatic highlighting via word order); (b) a sequence of a noun subject and an adverb, another canonical structure; and (c) a sequence of an adjective and a noun, i.e. a modified NP in a subject position. All of the three types of constituents occur at the beginning of a sentence since the pitch range in this position is the widest and hence tonal properties are more salient on the pitch tracks. Here are the examples of the three types of contexts.

The following figures of the pitch tracks of two consecutive words of this type show that an additional L tone is present between the two disyllabic words with a rising and a falling accent. This is unexpected given the theory of I&Z.

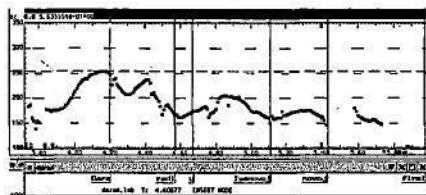


Figure 16: Noun-verb sequence. The sentence is *Dara radi u Ivanovoj novoj firmi.* 'Dara works in Ivanov's new company.'

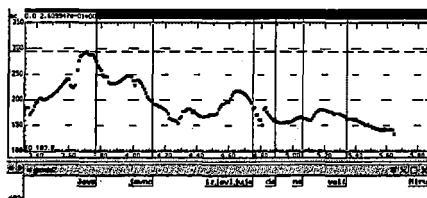


Figure 17: Noun-adverb sequence. The sentence is *Jóva jãvno izjavljuje da ne voli Miru*. 'Jova publically claims that he doesn't like Mira.'

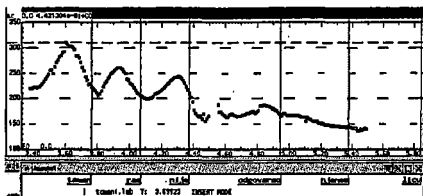


Figure 18: Adjective-noun sequence. The sentence is *Tãman rãm nije odgovarao njenom licu*. 'A dark frame didn't suit her face.'

In all three pitch tracks (Figures 16–18), the first word is disyllabic and has a long rising accent on the first syllable and the second word has a long falling accent on the first syllable. Since the first word is disyllabic, we know that the H tone will be realized on the second (i.e. final) syllable. The second word, having the falling accent on the first syllable must exhibit a H tone on the first syllable. If there were no word boundary tones, simple concatenation of these two words should produce a steady pitch line representing the two H tones, one from the final syllable of the first word and one from the initial syllable of the second word.

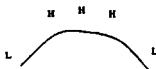


Figure 19: Schematic representation of the prediction for two consecutive H tones according to I&Z's theory.

SERBO-CROATIAN INTONATION

However, as we see in Figures 16–18, the two H tones are separated by a dip in pitch. This intervening valley I take to be the evidence for the %L boundary tone.

We may ask where the %L tone belongs. Does it belong at the end of a word? or is it the initial leading tone of every word, i.e., the beginning of every phonological word? For reasons that have to do with overall declination pattern, and patterns in sentence initial and final positions, I will assume that the word boundary tone is at the beginning of the word. I will argue for this hypothesis in section 4.2.2, where I discuss utterance final position.

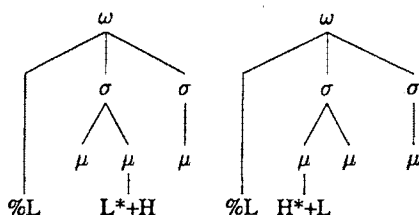
The acceptance of the L word boundary tone then gives us the following picture of the word tones for trisyllabic words with an accent on the first syllable.

	FALLING	RISING
SHORT	$ \begin{array}{c} \sigma \quad \sigma \quad \sigma \\ \quad \quad \\ \mu \quad \mu \quad \mu \\ \\ \%L \quad H^{*}+L \end{array} $	$ \begin{array}{c} \sigma \quad \sigma \quad \sigma \\ \quad \quad \\ \mu \quad \mu \quad \mu \\ \\ \%L \quad L^{*}+H \end{array} $
LONG	$ \begin{array}{c} \sigma \quad \sigma \quad \sigma \\ / \backslash \quad \quad \\ \mu \quad \mu \quad \mu \quad \mu \\ \\ \%L \quad H^{*}+L \end{array} $	$ \begin{array}{c} \sigma \quad \sigma \quad \sigma \\ / \backslash \quad \quad \\ \mu \quad \mu \quad \mu \quad \mu \\ \\ \%L \quad L^{*}+H \end{array} $

Table 10: Surface representation of tones in trisyllabic words with the stress on the first syllable and including the initial word boundary tone.

In terms of the theory of tone association to the prosodic hierarchy proposed in Pierrehumbert and Beckman (1988), the boundary tone is associated to the word node, whereas the lexical tones are associated to the stressed syllable. So, the tone structure of the sequence of the first two consecutive words depicted in Figures 16–18 can be represented as follows:

(2)



We are now ready to get back to the question we left at the end of the previous section regarding a possible explanation for the distribution of the rising accents. To review, we said that assuming that the accents in SC are bitonal, where only one of the tones is associated to the stressed syllable, the model requires no extra mechanism to explain the occurrence of the falling accents in monosyllables, as I&Z's theory would certainly require. The question that we could not answer at the time concerned the curious distribution of the rising accents: they never appear on monosyllables or on the last syllable in a word. Positing a word boundary tone at the beginning of the word creates the following sequence for the rising accents: $\%LL^*+H$. The $\%L$ tone serves the delimitative function, whereas the L^*+H (the pitch accent) serves the culminative and the contrastive function — the accent is rising not falling. The sequences $\%LL^*+H$, $\%LH^*+L$, and two durations would be hard to contrast on a single syllable. That is, I propose that it is the initial word boundary tone which creates an impossible sequence for monosyllables under the rising accents due to crowding of tones of the same type (i.e. L tone) with different functions, particularly when the duration of the starred tone needs to separate long accents from short ones. If rising accents did occur on monosyllables then we would need to be able to make a four-way distinction in the timing of the rise on a single syllable. At this point this is a very speculative statement and more research would be needed to confirm this hypothesis.

It is worth pointing out a historical perspective on the distribution of accents. The synchronic situation is a product of the so-called Neoštokavian stress shift (started in the 15th century). There were only the two falling accents in the old Štokavian dialects. The retraction of the stress from the syllable associated with the H tone to the preceding syllable gave rise to the rising accents. In other words, the rising accents are the reanalysis of the situation that arose when the stressed syllable was no longer associated with the H tone. This separation of the link between a stressed syllable and a H tone thus seems to be adequately captured in the proposal given in this paper.

4.1.2 Double-Accented Words

An additional piece of evidence for the $\%L$ word boundary can be found in double-accented words. The concept of a doubly accented word may seem odd since I am assuming

SERBO-CROATIAN INTONATION

that a definition of a phonological word is a prosodic unit with only one pitch-accent and, as I am also arguing here, a word boundary tone. However, there seem to be exceptions to my definition of phonological word. It is possible to find examples of double-accented phonological words. These words are always polymorphemic, and are in free variation with variants realized with one pitch accent. They give us an example of what a string of pitch accents looks like without a word boundary.

The following pitch track shows two near-identical sentences containing a word *najmanja* 'the smallest', which can have either the long-falling accent on the first syllable *naj-* or it can have two long-falling accents on the first and on the second syllable.³ The utterance on the left side contains the one-accented version and the utterance on the right side, the two-accented version.

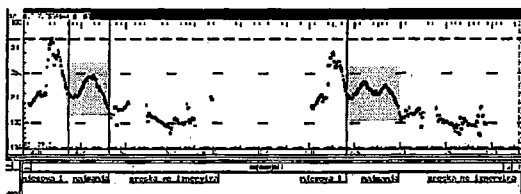


Figure 20: Two utterances of the sentence: *Njegova i najmanja greška me iznervira* '(Even) his smallest mistake irritates me.' The shaded parts of the f₀ represents the word *najmanja* in the two utterances; the one on the right contains a double-accented word.

If there were a word boundary tone in the double-accented word, then the L tone between the two peaks would have been lower and the two peaks would not have been of the same height (as is the case for succeeding words, since downstep is a part of every intonational phrase, and will be discussed in section 4.2.3). To see the difference between a double accented word and two words under the same accent as the double accented word, I provide the following pitch track, where the word *najmanja* occurs in the first utterance and the words *moj mali*, which have the same accent pattern across the same number of syllables, are in the second utterance.

³Stevanović (1989:431) notes that some long forms of superlatives obligatorily have two accents, such as *najdostojanstveniji* 'the most dignified'.

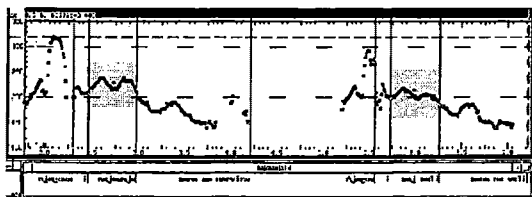


Figure 21: The first utterance, on the left, *Njegova i najmanja greška me iznervira* '(Even) his smallest mistake irritates me.' contains a double accented word; the second utterance, on the right, *Njegov i moj mali mene ne voli* 'His and my young one doesn't like me.' contains two words, with the same number of syllables as the first utterance, in the position of the double accented word. The shaded parts of the f0 point to the parts in the two utterances relevant for the comparison.

In a purely autosegmental account, this state of affairs is difficult to account for since all surface tones are associated to the tone-bearing units. In an autosegmental/metrical account, argued for here, the tone string is being decomposed into culminative and demarcative tones, which in turn are associated to different units in a prosodic structure.

To summarize, I have introduced a new concept into the description of the SC prosody, the %L word boundary tone. The evidence presented so far for the word boundary comes from two sources: the pitch dip observed in sequence of words under a rising and a falling accent, and the pitch level differences observed when this dip is compared to the dip in sequences of falling accents in polymorphemic words such as *najmanja*. These differences in F_0 pattern within a morphological word and across two words is easily explained in autosegmental/metrical account. The two pitch contours can be given two different parses by having the two L tones be part of different constituents in prosodic structure. In more classical autosegmental accounts with only one type of tone-bearing unit, on the other hand, both strings are analyzed as the same HLH sequence.

The third piece of evidence for a L word boundary tone will be introduced in section 4.2.3, where I will try to argue for the downstep model of the downtrend in SC. If the idea that downstep is a consequence of the alternation between H and L tones is correct, as suggested in autosegmental literature on African tone languages (see e.g. Clements and Ford 1979, 1981) then we might expect SC to use the H L alternation as a trigger for downstep regardless of where the L comes from in the grammar of tone. But, as can already be seen in Figure 20, the sequence of two peaks in double-accented word has the peaks at the same level, whereas the sequence of two peaks in two consecutive words, which I claim are separated by a %L word boundary tone, the two peaks are not at the same level. Consequently, it seems reasonable to speculate that it is the presence of a %L

word boundary that would account for the downstep model in SC very naturally, rather than saying that downstep occurs with any H L sequence. But before we can accept that as evidence, the nature of the downtrend needs to be examined in more detail to see whether a downstep account is tenable.

In this section, I also hypothesized that the presence of the %L boundary tone allows us to explain the distribution of the rising accents by assuming that the sequence of %LL* tones followed by +H needs more space for realization than a single syllable.

4.1.3 Clitics

Serbo-Croatian has morphological words which lack both stress and accent and thus are called toneless words, i.e., clitics. These words are prosodically dependent on phonological words. There are two types of clitics in SC, proclitics and enclitics. Prepositions are proclitics, they cliticize to the noun that follows them.⁴ Short forms of personal pronouns and verbal auxiliaries are enclitics. They are the so-called second position clitics, they cliticize to the preceding word. In this section, I show that proclitics and enclitics differ not only with respect to whether they precede or follow their host but also in their tonal specification. I argue that proclitics realize an edge tone, whereas enclitics have no tonal properties.

The shaded parts in Figures 22–24 show the prosodic behavior of a preposition (*prema* 'towards'), which is a proclitic, in three different positions in the sentence, initial, early medial, and late medial positions. Absolute final position of a preposition is not possible, since preposition stranding is not a syntactic option.

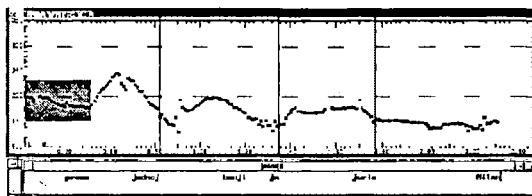


Figure 22: *Prema jednoj banji je jurio Milan* 'Milan was rushing towards the iodine spa.'

⁴Negative particle *ne* is also a proclitic, however I will exclude it from consideration in this paper. It cliticizes on to the verb that it modifies. Sometimes it even incorporates into the verb, i.e., *nisam < ne jesam* 'am.not'. When unincorporated, it can sometimes attract the accent *né znam* 'I don't know'. Prepositions do not attract the accent. It thus differs from prepositions, in the dialect I am describing.

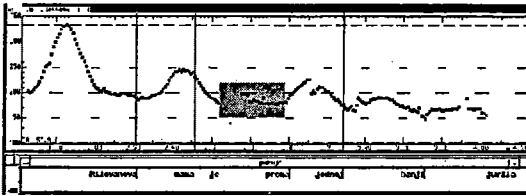


Figure 23: *Milovanova mama je prema jednoj banji jurila* 'Miľovan's mother was rushing towards the iodine spa.'

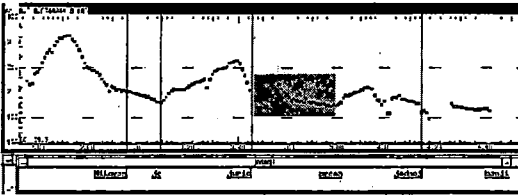


Figure 24: *Milovan je jurio prema jednoj banji* 'Milovan was rushing towards the iodine spa.'

The sequence of a proclitic and its host, *prema jednoj* 'towards iodine', in Figures 22–24, according to a traditional wisdom is a sequence of a toneless word and a word with the long-falling accent. In an autosegmental account, these moras would be assigned a L tone by default. In the autosegmental/metrical account, the proclitic would be realizing the left word boundary tone and the prediction would be that the F_0 associated with the preposition would be an interpolation between the edge tone and the accent.

However, the above three figures allow us to see that the F_0 of proclitics in all three positions is comparatively low and flat, and does not contain a peak. Moreover, the rise to the peak of the falling accent does not start until the beginning of the word that bears that accent. That is, the F_0 stretch relating to the preposition seem clearly separated from the F_0 relating to the host of the preposition. I propose that we analyze proclitics as a sequence that realizes a left edge tone, an initial %L word boundary tone. That is, proclitics add edges with no heads. Thus, a sequence of a proclitic and its host is a realization of two edge tones and an accent: %L %L T*+T. The motivation for this analysis comes from the F_0 on the proclitic, which starts low and stays low (or even falls slightly) until the beginning of the word the clitic is attached to. Since, in examples like above, we have a word under the long

falling accent, i.e. H*+L, without a %L at the host's edge we would predict a steady rise towards the accent H tone from the left edge of the proclitic. However, we always get a steep rise only at the beginning of the left edge of the host of the proclitic and not from the left edge of the proclitic itself.

In contrast to proclitics, enclitics do not have an edge tone associated with them. They are truly toneless morphemes. In SC, enclitics cluster in the so-called second position. The second position is an elusive concept because its best definition is a disjunction: 'the second position is either after the first accented word, or after the first accented constituent' (see Browne 1967:5, who was the first to discuss SC enclitic placement in the generative literature).

In Figures 22–24, we have an auxiliary clitic *je* occurring in various positions. In Figure 22 it occurs after the third phonological word, in Figure 23 after the second, and in Figure 24 after the first. In all of these figures we can observe that the clitic functions as material that interpolates between two tonal specifications: the accent of its host and the %L word boundary tone of the succeeding word.

To summarize, proclitics and enclitics differ prosodically. Proclitics carry a word edge tone, whereas clitics do not.

In this section I have argued, on the basis of tonal evidence, for a prosodic unit which I call the phonological word. This prosodic domain is defined by a word boundary tone as a delimitative marker and a pitch accent as a culminative marker. I have also shown that in some cases we have a unit which may lack a culminative marker, such as proclitics, or a unit which may lack a delimitative marker, such as double-accented words. These prosodic units are fused with other units that complement them to form a phonological word. It would be interesting to see if tonal evidence for this prosodic unit can be strengthened by segmental evidence as well.

4.2 Intonational Phrase

In this section I discuss a prosodic constituent higher than the phonological word, namely, the intonational phrase. Two major properties of this prosodic constituent are phrase accents, boundary tones, and pitch range manipulation. That is, pitch range expansion and contraction, and boundary tones can be used as probe for prosodic structure above the word in SC. I show that this prosodic constituent realizes four types of tones: two boundary tones, L% and H% in combination with two phrase accents, L- and H-.

4.2.1 Initial Position

Both the sentence initial position and the discourse initial position in an utterance have the highest H target of all the phonological words in a sentence. However, the two

differ by the level of H. The utterance initial H is higher than the sentence initial H. This position is set off from the rest of the words by the relatively higher pitch target regardless of the syntactic status of the constituent or the word. That is, the H tone in the first position is higher than the H in the second position regardless of whether the phonological word is a syntactic unit by itself or a part of a larger phrase.

To illustrate this point, consider a more elaborate utterance (in Figure 25) consisting of three sentences instead of just one. We can notice that the H in each subsequent sentence initial position is slightly lower than the preceding one. Thus, the absolute utterance-initial position is always set off from all the others by its highest H target. We can see this clearly in the pitch track in Figure 23, representing the following text:

- (3) a. Milovanova mama je žurila na voz.
 Molovan's mother aux hurried on train
 Milovan's mother was rushing to catch a train.
- b. Nije imala vremena da gleda ljude u prolazu,
 not.aux had time that look.at people in transit
 She didn't have the time to observe people around her,
- c. ali je njenu pažnju Marija ipak privukla.
 but aux her attention Mary still attracted
 but Mary still managed to attract her attention.

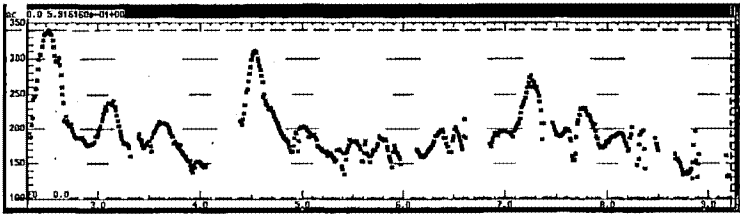


Figure 25: Three consecutive sentences from example (3), showing the set off peaks of the initial constituents and scaling of the three peaks across discourse.

Each pair of adjacent sentences in the above sequence is separated by a short pause, yet their initial H targets create an internal slope thereby bringing cohesiveness to the whole utterance. The internal structure of the three sentence utterance is reminiscent of English utterances as documented by Lehiste (1975). Lehiste showed that in English paragraphs, utterances are characterized by a certain intonation structure, the so-called 'paragraph intonation'. The relationship between pitch range and discourse topic structure has also been

suggested by Brown et al. (1980) and Hirschberg & Pierrehumbert (1986) for English, and by Grønnum (1985) for Danish.

4.2.2 Declarative Sentence Final Position

The sentence final position is also characterized by its distinctive intonational shape. Any type of a syntactic constituent with any type of a word accent in a sentence final position shows a highly reduced pitch range with the pitch very close to the speaker's base line. This effect is treated as final lowering in Inkelas and Zec (1988:240) or laryngealization by Lehiste and Ivić (1986:186). L&I point out that the effects of laryngealization very often seem to lead to neutralization of the accents in a sentence final position. This leads I&Z to posit the rule of final lowering, which stipulates the insertion of a L tone on the last syllable of the last word over-riding the H of the lexical accent (which in their model is always an associated tone). This rule makes a prediction that accents in disyllabic words are neutralized in sentence final position. The data that I have collected show that the distinctions among the word accents are still preserved (Godjevac 1999). However, the distinctions are reduced relative to the initial or medial positions in a sentence of this type. Hence, I would argue that a phonological representation should not include a rule like I&Z's final lowering, since the phenomenon appears to be an effect of some aspect of backdrop pitch range, which Figure 25 shows can be varied in continuous but systematic way to gradiently signal a position within the larger discourse.

The following two figures show the difference induced by the sentence position on the same words. In the first figure, Figure 26, we can see the initial position of the word *mlâda* 'young' and the final position occupied by the other member of this minimal pair, the word *mlâda* 'bride'. In the second figure, Figure 27, the two words are in the reversed positions. This illustration allows us to see the difference between a falling accent and a rising accent in the sentence initial vs. final position.

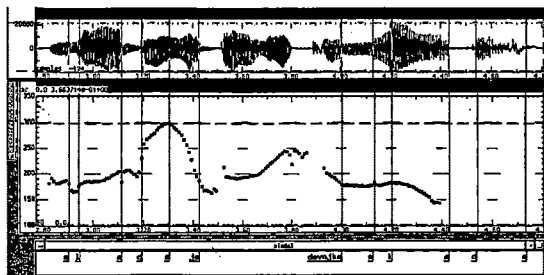


Figure 26: *Mlâda je devojka mlâda* 'A/The young girl is a/the bride.'

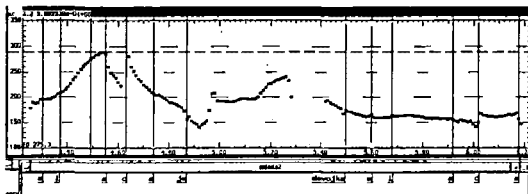


Figure 27: *Mlada je devojka mlada* 'A/The bride is a young girl.'

From the above figures we can see that the rising accent stays level in the final position, whereas the falling accent is falling, and it actually becomes laryngealized. Laryngealization is a low pitch common to final falling intonation. Note that in Figure 26 vocalization seen in the wave form continued past the point where the pitch extraction algorithm gives up. The wave form shows the irregular (laryngealized) pulses whereas the pitch track is empty. Therefore, there is a clear differentiation between the two accents even in the sentence final position. The reduction of the pitch range did not erase the lexical tonal distinctions.⁵

The rule of final lowering of I&Z is an insertion of a L tone on the final mora of an utterance. This rule predicts that the final syllable of *mlada* should be lower than the last mora of the first syllable, which would be assigned the lexical H. As we can see from the Figure 27, that prediction is not borne out.

Instead of positing a final L insertion rule, which effectively erases the lexical H, I posit a L- phrase accent. That is, declarative utterances are marked by a L- phrase accent, followed by a L% boundary tone. The accent and the boundary tone are properties of a higher level phonological constituent, the intonational phrase. They are realized by lowering the pitch range of the constituent that carries the phrasal marking: the right-most constituent in neutral prosodic conditions, or whatever constituent is chosen in the case of prosodic focus, as we will see in section 4.2.4.

The consequence of the higher level tones on the final position in neutral prosodic contexts can be observed by looking how it affects the final peak in longer utterances. It is clear that the peak in the final constituent's F_0 is lower than the proportional reduction

⁵There is some additional evidence for the preservation of the falling/rising distinction. In her acquisition study of SC accents, Kariya (1983:60) notes that 'the distinction between rising and falling accents was evident from patterns of post-stressed syllable deletion: the vowel in a syllable immediately after a falling accent was much more likely to be whispered or deleted than the vowel in a syllable immediately after a rising accent.'

SERBO-CROATIAN INTONATION

based on the preceeding peaks would have predicted. Schematically, we could represent that relationship in the following way:

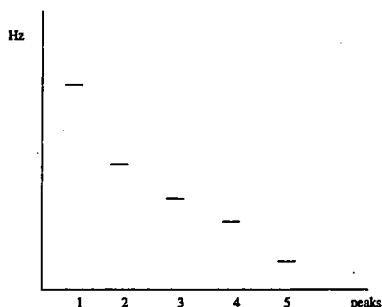


Figure 28: Schematic representation of the peak-proportions for an utterance of 5 phonological words.

I claim that this is a direct consequence of the two final L tones, L-L%, associated with the higher level prosodic constituent, the intonational phrase.⁶

Thus, what seemed like a conspiracy against lexical accents in final position is just a consequence of tonal marking of higher level prosodic constituents. Lexical accents are still present in the final position, but they are affected by the higher level tones. This analysis predicts that the shorter the content word in the final position, the more crowded tones will be, and consequently the more difficult it would be to see them by observing (measuring) the F_0 . Hence, under the assumption that tones are only properties of syllables, the conclusion that the accents are neutralized in this position seemed inevitable.

There is another piece of evidence that accents are not neutralized in the final position: they show up clearly under prosodic narrow focus. I will present this evidence in section 4.2.4, as a part of the discussion of prosodic focus.

Before I close this section I want to bring up again the question of where the L word boundary tone belongs. I proposed earlier that we assume that the word boundary tone belongs to the left edge, that is, at the beginning of every word. My reasoning for this has to do with the intonational phrase initial words. Since falling accents are specified for a

⁶This property may be similar to what Liberman & Pierrehumbert (1984) found for English and for which they proposed a phonetic rule of final lowering. If under a more scrupulous investigation the sequence of the two L tones cannot account for the F_0 in the final position, a rule of final lowering analogous to the rule proposed for English would also be necessary for Serbo-Croatian.

HL melody we would expect that they would start higher than words with rising accents which have a LH melody. However, if we compare the initial constituents in Figure 26 and Figure 27 (which are minimal pairs), we see that both start with the same F_0 , which is relatively low. The fundamental frequency of the word in Figure 26, which is under a rising accent, stays low, whereas the word in Figure 27, which is under a falling accent, rises steeply to reach its H tone. It is easy to account for this similarity in the F_0 pattern of the beginnings of words, if we postulate a L word boundary tone at the left edge of a word.

4.2.3 Downtrend

From the all figures presented thus far we can also observe that the pitch contour of the SC declarative, broad focus utterance exhibits a certain downward motion. That is, each subsequent phonological word in the sentence has a lower H target than the preceding one. This behavior of the declarative tune needs to be accounted for. The decline in the pitch level as a declarative utterance evolves seems to be a fairly common phenomenon crosslinguistically (Ladd 1996:73).

Modelling of the pitch decline across an utterance can be done in several ways. One model reduces the high and the low tones in a declination mode equally, keeping the tonal space the same over time. A different model reduces high tones (Pierrehumbert & Beckman 1988). A more complicated model reduces both high and low tones but each of them differently (Pierrehumbert 1980). A schematic representation of these models is illustrated in the following figure.

SERBO-CROATIAN INTONATION

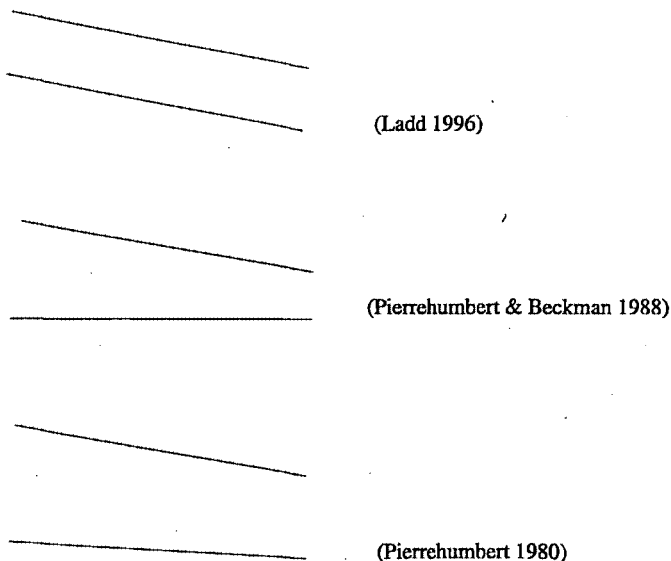
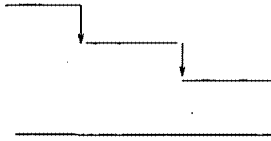
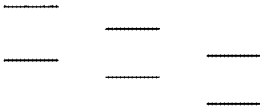


Figure 29: Declination models

However, declination is not the only way to account for the pitch decline across an utterance. Pierrehumbert (1980) has shown that for English, it is also possible to assume a downstep model. The difference between a downstep model and a declination model is in the predictions of the way pitch level is realized between relevant peaks. According to a declination model, the pitch level declines all the time, that is, even between the relevant peaks. According to a downstep model, the pitch level is level between the relevant peaks and it only declines in a step motion at a relevant point. Downstep models can also be modelled in several ways, similarly to the declination models mentioned above. A schematic illustration of the downstep models is presented in the following figure.



(Pierrehumbert 1980)

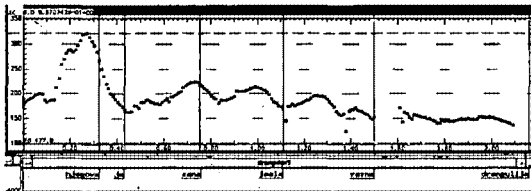


(Ladd 1996)

Figure 30: Downstep models

L&I have observed that Serbo-Croatian sentences show an overall downtrend. Their data offer a possible answer to the question regarding modelling of this downtrend. As they show in their figure 3.6, the peaks and valleys of falling and rising accents scale differently. The downtrend across peaks is steeper than that for the lows, and moreover, the peaks and valleys of the falling accents are higher and lower respectively, than the peaks and the valleys of the rising accents. The precise modelling of the downtrend is not possible with the current data. However, the informal evidence suggests that this downward slope is not a continuous declination of the pitch but rather a downstep of the highs and possibly lows within an intonational phrase. In addition to the above mentioned models of downtrend, it is possible to imagine a model that may involve a combination of a downstep of the highs and a declination of the lows. In what follows, I will show what kind of evidence we have and will tentatively argue for a downstep model, although precisely which type of a downstep model will be left open.

A typical effect we find in connection to downtrend in SC can be seen clearly in Figure 31.

Figure 31: *Njegova žena je imala razne drangulije* 'His wife had all sorts of junk.'

SERBO-CROATIAN INTONATION

In Figure 31, we can notice that the H of every word (i.e. the lexical H tone, which is part of every phonological word regardless of the accent) is slightly lower than the preceding one, modulo the first and the final words, which seem to be subject to their special position in an utterance. Thus, there is a clear effect of the downtrend in a declarative utterance. If we assume that the pitch range falls steadily throughout the utterance, as in a declination model, then this is not surprising. What a declination model also predicts is the steady decline of the pitch even in syllables that are marked for the same tone. So, a good testing ground for this prediction would be an utterance consisting of longer words whose accent is later in the word. A good candidate for this would be the word *omalovažavanje* 'humiliation'.

The following three figures show the word *omalovažavanje* 'humiliation' in the three sentence positions, initial, medial and final, respectively.

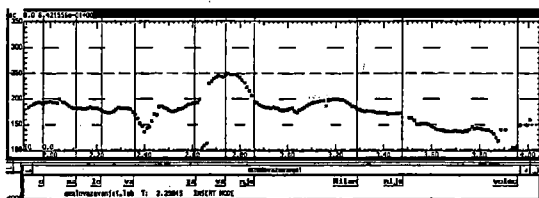


Figure 32: *Omalovažavanje Milan nije voleo* 'Humiliation, Milan didn't like.'

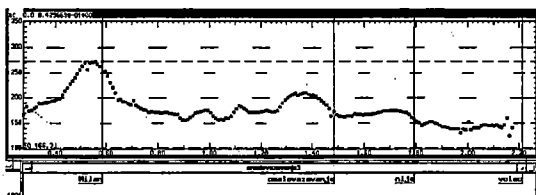


Figure 33: *Milan omalovažavanje nije voleo* 'Milan didn't like humiliation.'

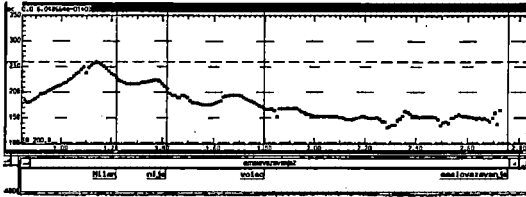


Figure 34: *Milan nije voleo omalovažavanje* 'Milan didn't like humiliation.'

The pitch tracks in Figures 32–34 show us that the syllables which are not affected by the lexical accent or the boundary tone in sentence initial and sentence final positions do not stay level completely, but seem to show a slight slope, whereas in medial position they stay level. This is even more prominent in Figure 35 in which the relevant portions are shaded:

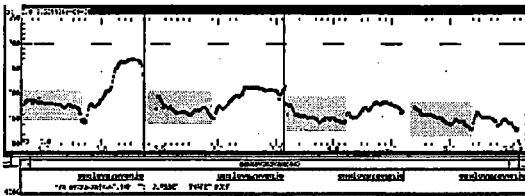


Figure 35: *Omalovažavanje, omalovažavanje, omalovažavanje, omalovažavanje* 'Humiliation, humiliation, humiliation, humiliation.'

Schematically, the slopes of the unaccented syllables found in Figure 36 (the shaded areas) can be represented in the following way:

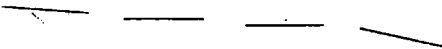


Figure 36: Schematic representation of slopes in Figure 35.

Figure 36 schematically represents the slight slope of the unaccented syllable strings in initial and final positions in a sentence. We do not observe the same effect in medial positions. Since a declination model would predict a slope in medial positions, I

SERBO-CROATIAN INTONATION

propose a downstep model as an account of the downtrend within an intonational phrase.⁷ Although these stretches may seem to be too short for a definite conclusion regarding the L tones, we may confidently say that L tones are not subject to whatever it is that is reducing the successive H tones to the same degree. Sentence initial position and sentence final position would have to be accounted separately. However, given that these two positions have additional properties not shared by others (discussed in section 4.2.1 & 4.2.2), they require a special treatment anyway.

I want to introduce another property of the SC prosodic system that I will call a 'pleating effect', which seems to be a direct function of the length of the utterance and is relevant for any modeling of the downtrend. To my knowledge, this was first discussed in Kostić (1983). Basically, the pitch range gets partially reset to a higher target at constituent boundaries as the utterance gets longer.⁸ This effect has also been noted for Japanese by Kubozono (1992), although he called it 'metrical boost' and gave it a specifically rhythmic interpretation. As he explains, the phenomenon:

'... can be understood [in such a way] that the downstepped phrase has been raised by the phonetic realization rule of metrical boost to such an extent that it is now realized higher than the [previous] phrase. This case is typical ... at major syntactic boundaries ...'

I will illustrate this phenomenon in SC by a series of three pitch tracks that represent a successive lengthening of a simple sentence. The three sentences are as follows:

- (4) a. Njegova žena je imala dve violine.
his.NOM wife.NOM AUX had two violins.ACC
'His wife had two violins.'
- b. Njegova žena je imala dve violine iz istog perioda.
his.NOM wife.NOM AUX had two violins.ACC from same period
'His wife had two violins from the same period.'
- c. Njegova žena iz prvog braka je imala dve violine iz istog perioda.
his.NOM wife.NOM from first marriage AUX had two violins.ACC from same period
'His wife from his first marriage had two violins from the same period.'

An utterance of the sentence in example (4a) has no pleating effect, as the following pitch track shows:

⁷The slight difference in the medial stretches defined by L tones can be explained by treating the L word-edge tone and the L* of the accent as different targets for the L.

⁸The partial reset of the declination was discussed in Ladd (1984, 1988); however the partial reset was a function of scope disambiguation between two conjunctions, 'and' and 'but'. The partial reset may have the same function in SC as well, but, it need not, as in the case I am presenting. It can simply be a function of the length.

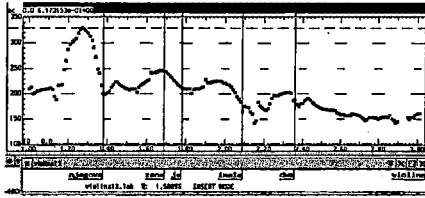


Figure 37: *Njegova žena je imala dve violine* 'His wife had two violins.'

As we can see, the H targets get lower and lower in the utterance as we proceed from the beginning to the end. The next two pitch tracks illustrate a 'pleating effect'.



Figure 38: *Njegova žena je imala dve violine iz istog perioda* 'His wife had two violins from the same period.'

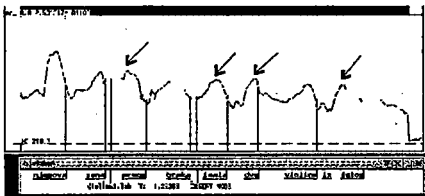


Figure 39: *Njegova žena iz prvog braka je imala dve violine iz istog perioda* 'His wife from his first marriage had two violins from the same period.'

The pitch range reset at each phrase is done in such a way that the level of the H tone is reset to the same level, or a slightly higher level than the preceding H tone, thereby

breaking up the downtrend. This is the 'pleating' effect. There is no focal prominence on any of these constituents in which the first H was reset to a higher pitch range.

Beckman & Pierrehumbert (1986) show, the reset of the pitch range can serve as evidence for a phrase boundary, as they used it, in addition to pausing, in arguing for an intermediate phrase in English. So, is the 'pleating effect' then a matter of prosodic phrasing? That is, do the points of reset correspond to prosodic boundaries of any sort? And if so, what type of prosodic constituent do these points correspond to? I propose that these units are phrases which are the domain for the local manipulation of the pitch range, i.e. downstep, as also proposed for English by Pierrehumbert (1980). I will call them intermediate phrases.

Formation of intermediate phrases is sensitive to syntactic boundaries. However, the syntactic boundaries that seem to be relevant do not form a natural class. Consider the resetting of the pitch range in Figures 38 and 39. In Figure 38, the reset is done at a boundary between a noun phrase and its PP modifier (adjunct). In Figure 39, there are four reset points: (1) at the same point as in Figure 38, (2) at the point of a syntactic head/complement boundary (V and NP), (3) at the point between the last constituent in the subject NP and the first constituent of the VP (i.e., the main verb), and (4) at the boundary between an NP and its PP modifier (the same boundary as in (1)). These are the two basic types of syntactic boundaries: head/complement and head/modifier. Because both types of syntactic boundaries can function as reset points, I take this to be evidence that intermediate phrases cannot be derived by an algorithm sensitive to syntactic relations, such as the one proposed by Nespor & Vogel (1986). In addition, an end-based algorithm, as proposed by Selkirk (1986), also does not make the correct prediction. An end-based algorithm would predict intermediate phrases in shorter utterances, such as those depicted in Figure 37, where we never find them.

As I have shown, and as Kostić (1983) has also claimed, intermediate phrase formation is a function of the length of an utterance. Kostić claims that the relevant crossover point is 5 words. That is, utterances that are longer than five words will inevitably be realized as more than one grouping of words, or in our terminology more than one intermediate phrase. However, how many intermediate phrases an utterance of six words will have is not determined. As Kostić argues, there could be two or three. That is, we expect speakers to differ in the way they chunk the utterance. Thus, even though syntactic boundaries are relevant to the formation of intermediate phrases, knowing where the syntactic boundaries are will not necessarily give us the correct grouping of words into intermediate phrases, because they differ both within speakers and across speakers.

4.2.4 Prosodic Focus

So far, we have looked at utterances which do not have any prosodic prominence except for the lexical stress. That is, prosodically they are all broad focus. Semantically, however, the focus domain is determined by the interaction of this prosodic property and word order considerations. So, prosodic broad focus is what I call neutral intonation.

In this section I turn to prosodic focus. By prosodic focus, I mean prosodically marked emphasis on some constituent in a sentence. Serbo-Croatian allows its constituents to be prosodically focused, which in turn signals semantic focus as well. Since semantic focus is crucial for interpretation of utterances, both for their truth conditional and non-truthconditional meaning, the investigation of prosodic focus is crucial in the overall understanding of the language. Semantic focus in SC can be signaled via word order as well as prosodically. For word order to function as semantic focus marking, prosodic focus must be absent. That is, the sentence intonation must be neutral.

Any phonological word (words that can bear accent) can be prosodically focused regardless of its position in the sentence and its syntactic function. The phonetic effects of prosodic focusing are pitch range manipulation of the focal constituent and its environment. A focal constituent is realized in a slightly expanded pitch range, whereas post-focal constituents are realized in a significantly reduced pitch range. In addition, pre-focal constituents may also be affected by a slight compression of the pitch range. The following five figures show the same sentence with different prosodic focus patterns. The first figure shows the sentence *Jelena daje Mariji limun* 'Jelena is giving Mary a lemon' in a broad focus utterance. The next four figures show the same sentence with a prosodic narrow focus on one of the constituents in the sentence, a different one in each case.

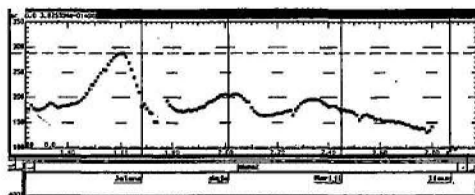


Figure 40: *Jelena daje Mariji limun* 'Jelena is giving Mary a lemon.'

SERBO-CROATIAN INTONATION

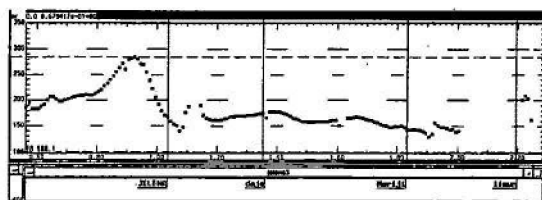


Figure 41: *Jelena daje Mariji limun* 'JELENA is giving Mary a lemon.'

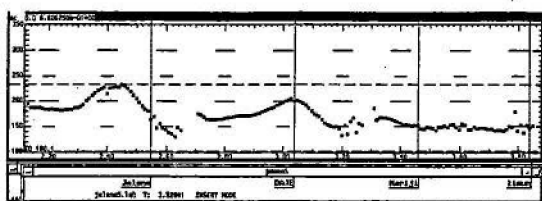


Figure 42: *Jelena daje Mariji limun* 'Jelena is GIVING Mary a lemon.'

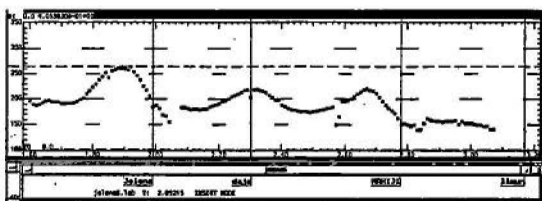


Figure 43: *Jelena daje Mariji limun* 'Jelena is giving MARY a lemon.'

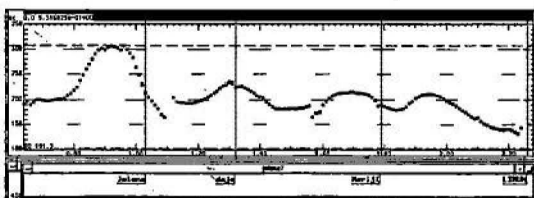


Figure 44: *Jelena daje Mariji limun* 'Jelena is giving Mary a LEMON.'

Looking at the peaks, we may notice that the prosodic focus on *Jelena*, in Figure 41, didn't raise the peak of this word but instead has reduced the peaks of the subsequent constituents. Prosodic focus on the verb seems to have lowered the H of the initial word and also reduced the peaks on the subsequent words. In Figures 43 and 44, the peaks of the focused constituent seem to be higher than in the broad focus utterance, Figure 40. We can also look at what happens to the final constituent. There seem to be three types of realization of this word: in a broad focus utterance, Figure 40, after a prosodic focus, Figures 41, 42, and 43, and being prosodically focused itself, Figure 44. A broad focus utterance gives the final constituent a reduced pitch range. The constituents following prosodically focused constituent manifest a much flatter pitch line. In other words, narrow focus affects post-focal constituents via pitch range reduction. The pitch range manipulation can be represented in the following way:

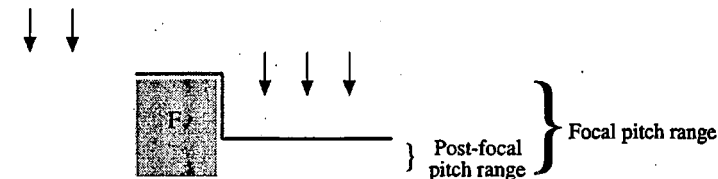


Figure 45: Schematic representation of pre-focal and post-focal pitch range reduction.

Prosodic focus affects the final constituent by widening the pitch range for this constituent, which is the reverse of what is observed for this position when it is not prosodically focused. This expansion of the pitch range for the final constituent allows the manifestation of the lexical accent with no reduction. This is another piece of evidence that the final position does not neutralize the accents (see section 4.2.2).

The next figure shows a familiar utterance from section 4.2.3 with the narrow focus on *žena* 'woman'. Being a longer utterance, the effect of prosodic focus is more obvious in the pitch track of this utterance than in a shorter utterance.

SERBO-CROATIAN INTONATION

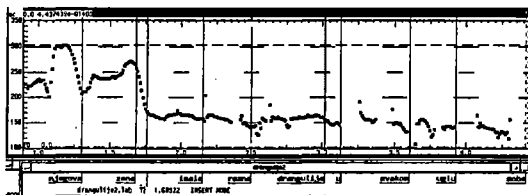


Figure 46: Another illustration of the pitch range reduction after the prosodic focus:
Njegova žena je imala razne drangulije u svakom uglu sobe 'His WIFE had all sorts
of junk in every corner of the room.'

The same effect of prosodic focus is reported for Mandarin Chinese (Jin 1996) and for Hindi (Harnsberger & Judge 1996). Jin shows that post-focused constituents (post-stressed syllables in his terminology) are affected by a significant pitch range reduction, whereas pre-focused constituents are not. According to Harnsberger & Judge (1996), Hindi also signals prosodic focus by drastically reducing the pitch range of the post-focal constituents, a phenomenon which they call register compression.

How can we account for the pitch range effect due to prosodic focus? I propose that focus is signaled by a phrase accent. The L- tone of the phrase accent is realized at the right edge of the word which is focused. That is, the phrase accent is realized earlier than the right edge of its phrase and spans over the string in the post-focal domain, which in turn lowers the pitch range for those constituents.

4.2.5 Morphologically Marked Questions

In this section, I look at the intonation of three types of questions: two types of yes-no questions, both of which employ the question particle *li*, and standard *wh*-questions.

The point of this section is to show that there are no prosodic differences between declaratives (that we have looked at thus far) and morphologically marked questions. That is, there is no special intonation necessary if the interrogative mood is morphologically specified. I look at *wh*-question, and yes-no questions.

Unless they are echo-questions, *wh*-questions obligatorily have the *wh*-word at the beginning of the sentence. In syntactic terms, *wh*-movement is obligatory. Grammatical status of the *wh*-constituent, argument vs. adjunct, does not affect the prosody of questions. Since *wh*-words are clause initial, their prosodic pattern is of the sentence initial position, discussed in 4.2.1, as the following pitch tracks show:

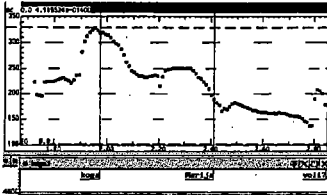


Figure 47: *Koga Marija voli?* 'Who does Mary love?'

Wh-words can also be focused, in which case the prosodic focus effects are the same as in declaratives; the post-focal constituents are in a drastically reduced pitch range. Compare the declarative prosodic focus, Figure 46, and the prosodic focus on the *wh*-word found in the following figure.

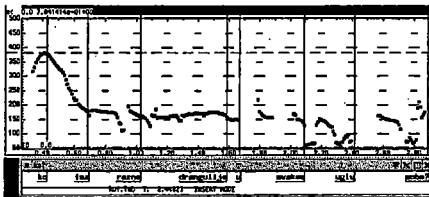


Figure 48: *Ko ima razne drangulije u svakom uglu sobe?* 'Who has all sorts of junk in every corner of the room?'

Yes-no questions are formed in several ways. The standard way is to start the question with *da li* (*Da li Marija voli Milana?* 'Does Mary love Milan?'); *li* is a question particle, and *da* is a complementizer 'that'. Another way is to start the question with *je li* (*Je li Marija voli Milana?* 'Does Mary love Milan?'), *je* is the short (clitic) form of the 3p.sg.pres. of the verb 'to be'. Clitics are by definition unaccented forms; however, when *je* precedes the question particle, it bears a short-falling accent.⁹ And finally, the third way is to attach *li* to the tensed verb or some other constituent that is being questioned. For the purpose of illustration, here are some examples of the third strategy:

- (5) a. *Milan li je otišao?*
 Milan.NOM li AUX left
 'Was it Milan that left?'

⁹That is, in order to support a clitic it must be prosodically "promoted" to an accented form.

SERBO-CROATIAN INTONATION

- b. Ode li Milan?
left.AOR li Milan
'Did Milan leave?'
- c. Kući li je Milan otišao ?
home li AUX Milan left
'Was it home that Milan left?'

A pitch track of a standard *da li* question is no different from a simple declarative utterance, as the following figure shows:

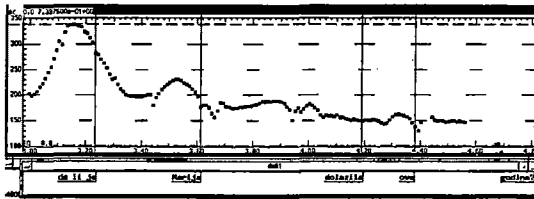


Figure 49: *Da li je Marija dolazila ove godine?* 'Did Mary come this year?'

However, the second type of *yes-no* questions, those with *je li*, seem to favor some additional focal prominence, most particularly on the verb, as also noted by Lehiste and Ivić 1977. Prominence on the verb is also found in Russian morphologically unmarked questions (Ladd 1996). For example:

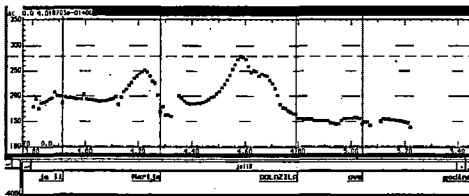


Figure 50: *Je li Marija dolazila ove godine?* 'Did Mary COME this year?'

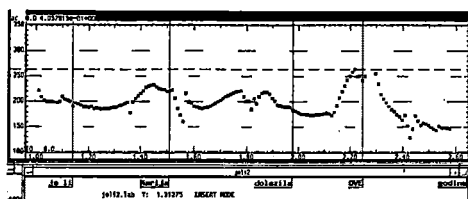


Figure 51: *Je li Marija dolazila ove godine?* 'Did Mary come THIS year?'

Yes-no questions create high H targets for the lexical H tones of the focused constituents, higher than prosodic focus in declarative sentences seems to produce (compare Figure 42 with Figure 50). There is a difference between a *da li*-question and other types of morphologically marked yes-no questions. *Da li*-questions have the initial high rise, just like declaratives.

To summarize, common to all questions is the fact that the final constituents do not exemplify a rise intonation. That is, there is no H% boundary tone at the end of a morphologically marked question. However, as we will see in the next section on question tags and the section on prompting intonation, the H% tone can mark utterances as questions when they are not morphologically marked (just as in English).

4.2.6 Question Tags

Another way to ask a question, employing morphology, is to use a question tag *zar ne?* or *jel' da?* 'isn't it the case?'. The basic contour of these types of questions involves a rising intonation at the end. I use these utterances as evidence for a H% boundary tone of an intonational phrase.

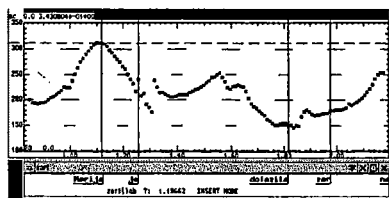


Figure 52: *Marija je dolazila, zar ne?* 'Mary came, didn't she?'

SERBO-CROATIAN INTONATION

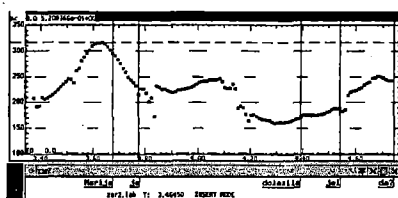


Figure 53: *Marija je dolazila, jel' da?* 'Mary came, didn't she?'

On the basis of contrast between the boundary tones in question tags and declarative utterances, I propose two different right edge intonational phrase boundary tones: L% and H%. Together with the phonological word left edge boundary tone, these tones are some of the markers of prosodic structure.

In the next section I introduce three new markers of prosodic structure: a L- phrase accent and H% boundary tone found in prompting intonation, a %H word boundary tone, found in double focus constructions, and H- phrase accent found in vocative chant. I discuss these contexts in a separate section because their tonal properties interact with lexical tonal specification that leads to loss of lexical information.

5 Loss of Lexical Information

5.1 Prompting Intonation

Prompting intonation can be characterized as the intonation pattern used for elicitation of information about some constituent. For example, it could be the intonation contour on *Marija?!?* which can then have the meaning of: 'What about Mary? Tell me something about her.' This intonation pattern can also be used for signaling a yes-no question, or for signaling surprise. L&I have studied this intonation pattern as a question intonation for morphologically unmarked yes-no questions. They name it 'a reverse pattern'. I will continue to call it prompting intonation in accordance with the terminology used by I&Z. As we will see, this intonation pattern seems to neutralize the lexical accents' patterns, the claim also made by L&I: 190. Prompting intonation then is an intonational morpheme that seems to overwrite the phonemic distinctions made by the lexical accents.

In a constituent under the prompting intonation there is a steep rise immediately after the stressed syllable. One hypothesis would be that this rise could be represented as a H% boundary tone. However, I will argue that prompting intonation is not just a simple H% boundary tone, but that it is a sequence of L- phrase accent followed by a H% boundary tone. The reason for this will become clear when we look at the pitch tracks of this contour.

In the following figures we see minimal pairs of the falling/rising opposition in prompting intonation. All four words have the stress on the first syllable.

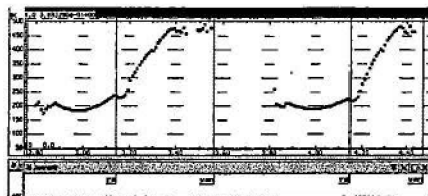


Figure 54: Long Falling/Rising accents: the minimal pair *ravan* 'plain' (falling accent) and *it ravan* 'flat' (rising accent) in prompting intonation.

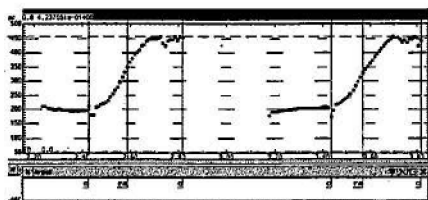


Figure 55: Short Falling/Rising accents: *dravo* 'he plowed' (falling accent) and *dravo* 'eagle' (rising accent) in prompting intonation.

The above figures show us that it is very hard to see any distinction among the minimal pairs in falling/rising opposition, as we saw it in declarative utterances. That is, the F_0 of the stressed syllable of the falling accents seems to be very similar (although there are some very small differences) to the stressed syllable of the rising accents in this intonational pattern even though according to their lexical specification we would expect them to be different, as they are in the declarative intonation pattern. Thus, this intonation pattern is a candidate for accent neutralization environment.

According to my data, and also according to L&I's analysis, all the accents seem to be neutralized under the prompting intonation in terms of their F_0 values.

We can see that prompting intonation affects the portion starting at the stressed syllable by looking at words with late accent placement, as illustrated in the pattern on the, by now familiar, word *omalovažavanje*, in Figure 56.

SERBO-CROATIAN INTONATION

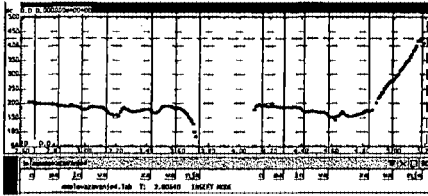


Figure 56: Long-rising accent: *omalovažavanje* 'humiliation' in citation form and prompting intonation.

The fact that preaccentual syllables are not affected by the prompting intonation allows words with late accents to be more easily distinguished, since only rising accents occur on non-initial syllables.

Monosyllabic words (which can only bear a falling accent) also show a pattern that is hard to account for if we assume that the lexical information is preserved under this intonation pattern, since there is no post-stressed syllable. The following pitch-tracks show the long-falling and the short-falling accent in a prompting intonation of a monosyllabic word.

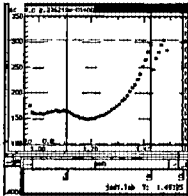


Figure 57: Long-falling accent
(*jod* 'iodine').

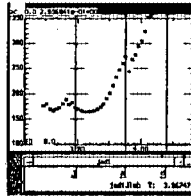


Figure 58: Short-falling accent
(*jäd* 'grief').

As Figures 57–58 show, the prompting intonation can also be realized on a single syllable. The basic pattern of this intonation type is preserved; the super-high target is realized in the second half of the syllable, even in the word under the short-falling accent, which is monomoraic, as discussed in section 3.3.4. These examples provide evidence that tones associated to structurally higher units can overwrite the tonal specification from lower levels.

To account for this intonation pattern I propose a sequence of L- phrase accent followed by a H% boundary tone. There is a difference however in the alignment of this

phrase accent and the L-phrase accent that we see in declarative utterances. The L tone of this phrasal accent is anchored to the stressed syllable of the last word (or the focused word – see Figures 62 and 63), rather than being realized over the metrically non-prominent ultimate (or sometimes ultimate and penultimate) syllable of the rightmost word. Grice et al. (in press) show that this is a characteristic of question accents in a number of unrelated Eastern European languages and their varieties, such as Hungarian, Romanian, and Greek. Serbo-Croatian has evidently also acquired this areal property.

Prompting intonation is also a prosodic focus marker, albeit with a question/surprise semantics rather than emphasis alone. Indicative sentences can be given interrogative mood with this intonation pattern. Figures 59 and 60 are examples of morphologically unmarked yes-no questions under the prompting intonation. Reversing the word order in the question produces a different focus, as indicated by the translation.

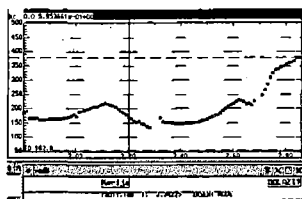


Figure 59: *Marija*
DOLAZI? 'Mary is
COMING?'

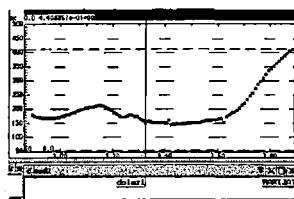


Figure 60: *Dolazi MAR-IJA?* 'MAY is coming?'

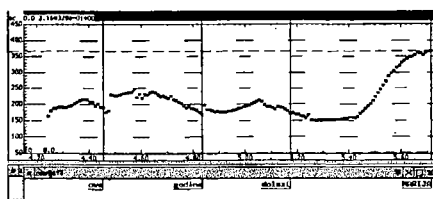


Figure 61: *Ove godine dolazi Marija?* 'MAY is coming this year?'

These examples are interesting because of the interaction between prompting intonation and focus of the question. The focus of the question is the word which bears the phrase accent, i.e. the edge constituent in the above examples. To a limited extent, it is

SERBO-CROATIAN INTONATION

possible to extend this edge. The following pitch tracks illustrate this point with both rising and falling accents.

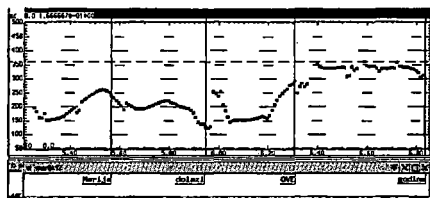


Figure 62: *Marija dolazi ove godine?* 'Mary is coming THIS year?'

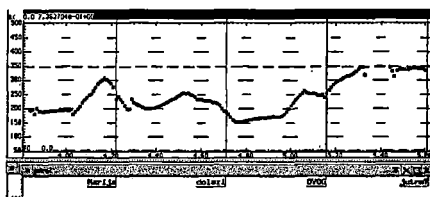


Figure 63: *Marija dolazi ovog jutra?* 'Mary is coming THIS morning?'

As we can see in Figures 62–63, the focused constituent is under the prompting intonation, and the constituent after it is in a highly raised and compressed pitch range. Falling/rising distinction also seems to be lost in this position. The length of the stretch following the H boundary tone seems to be limited to relatively short strings. As is even more clear in these examples, the L tone is anchored to the stressed syllable of the focused constituent, whereas the H tone is always at the edge. In other words, the two tones are timed differently.

As Ladd (1996) and Grice et al. (in press) show, in Hungarian, Romanian, and Greek questions are marked by the sequence L* H L, where the L* targets the stressed syllable of the focused word, and the HL sequence follows. In Hungarian the HL sequence targets the last two syllables of the phrase, whereas in Greek and Romanian the H tone of this sequence will target a stressed syllable if there are any. Thus, Serbo-Croatian prompting intonation differs from the one in the surrounding languages in the fact that Serbo-Croatian the tonal sequence is bitonal rather than tritonal, as it is in these languages. Serbo-Croatian does not have the final L boundary tone found in these languages.

5.2 Double Focus

In this section I introduce a new structural marker, a %H word boundary tone. I claim that this tonal marking occurs on certain focal constituents, such as the dependent variable in a double focus construction, or contrastive and metalinguistic focal prominence.

A double focus construction is an utterance which is an answer to a multiple wh-question, such as 'Who ate what?'. These types of utterances in English were first discussed by Jackendoff (1972) and later by Liberman and Pierrehumbert (1984). Jackendoff's example, 'FRED ate the BEANS', was described in terms of his A and B accents. The B accent is associated with the independent variable (the value for the wh-word that was fixed first) and the A accent with the dependent variable. This construction is also found in SC, and as I show creates contours very different from the ones we have already seen. Consider the following pitch track.

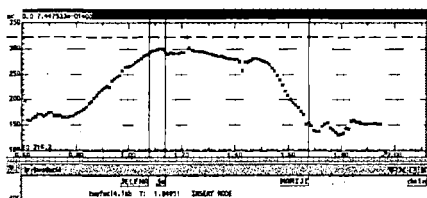


Figure 64: *Jelena je Mariji dala.* 'Jelena gave to Mary.' This utterance was an answer to the question *Ko je kome dao limun?* 'Who gave a lemon to whom?'

My analysis of the above contour is that the utterance consists of two intonational phrases: [*Jelena je*]_{IP1} [*Mariji dala*]_{IP2}. The first intonational phrase (IP1) contains a phrase which functions as the independent variable, and it is under prompting intonation. This tonal string then is %L L- H% (a word boundary tone and a L- phrase accent and H% boundary tone (prompting intonation)), where lexical tones are possibly overwritten. The IP2 starts with %H because the phonological word whose edge coincides with the left edge of the intonational phrase functions as the dependent variable. So, the tonal string in IP2 is %H L*+H L*+H L%, with pitch compression after the focused constituent. Evidence for this analysis comes from utterances where the two types of focus are separated.

SERBO-CROATIAN INTONATION

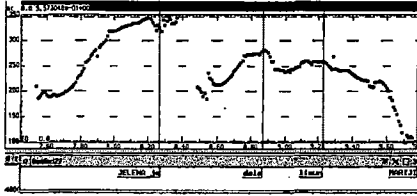


Figure 65: *Jelena je dala limun Mariji.* 'Jelena gave a lemon to Mary.' This utterance was an answer to the question *Ko je kome dao limun?* 'Who gave a lemon to whom?'

When the independent focus is not followed by the dependent focus, there is usually a break between the two phrases. In the above utterance, Figure 65, the dependent focus was placed at the end of the utterance. We can see that even in the final position, which as we have seen is low in broad focus utterances and especially low in non-focused utterances, there are two pieces of evidence for %H boundary tone: (a) there is no dip in the pitch contour signalling the L word boundary and, (b) signaling the finality of the phrase requires a much steeper fall, since the %H word boundary tone has raised the pitch range for the final constituent.

The %H word boundary tone affects the shape of the rising accents of the word to which it is attached, but the falling/rising opposition of accents is still distinguished, as shown by the following pitch track, which has a falling accent on the last word, as opposed to a rising accent in the previous utterance.

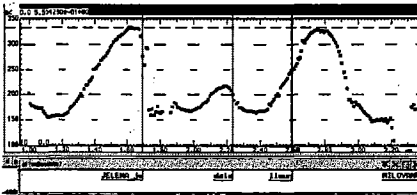


Figure 66: *Jelena je dala limun Milovanu.* 'Jelena gave a lemon to Milovanu.' This utterance was an answer to the question *Ko je kome dao limun?* 'Who gave a lemon to whom?'

However, the lexical accents preceding the %H tone seem to be affected. We can compare the F_0 shape of *limun* 'lemon' in the preceding figure with the F_0 shape of *ravan* 'flat one'

in Figure 67. The two words are similar enough for comparison, but differ in falling/rising opposition.

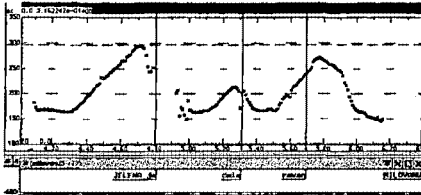


Figure 67: *Jelena je dala ravan Milovanu.* 'Jelena gave a/the flat one to Milovanu.' This utterance was an answer to the question *Ko je kome dao ravan?* 'Who gave a/the flat one to whom?'

Compare Figure 66 to Figure 67, which is a broad focus utterance of the same sentence.

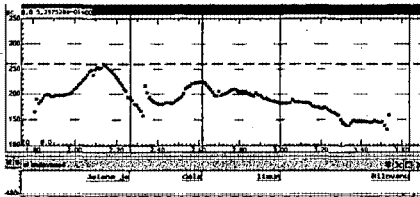


Figure 68: *Jelena je dala limun Milovanu.* 'Jelena gave a lemon to Milovanu.'

On this analysis the difference between the two pitchtracks consists in the phrasing, one intonational phrase in Figure 68 vs. two intonational phrases in Figure 66. In addition, the tonal strings are also different. In Figure 68, all lexical tones are preserved. In Figure 66, the lexical tones of the word 'lemon' are affected by the %H word boundary tone of the following focal constituent, as the comparison of the two figures clearly shows.

Thus, I conclude this section by noting that the %H word boundary tone affects the lexical information of preceding constituent. I now turn to the last section in which I discuss the vocative chant intonation.

5.3 Votive Chant

Votive chant is another intonational contour which seems to affect the lexical information to a great degree. The melody that characterizes votive chant is similar but not identical to the English votive chant. I&Z observe that 'the votive chant has a basic (Low)-High-Mid melody', where the Low is present only in words with three syllables or more and with the stress on non-initial syllable.

According to my findings, the initial L tone is present in all cases, which in this system is accounted by the %L word boundary tone. The rest of the shape of the F_0 contour shows a rise towards a H target and a continuation with a slight drop in pitch. This basic pattern shows up on all words regardless of their metrical structure. That is, what is common to all words under votive chant intonation pattern is the H tone on the penultimate syllable and a lower tone on the final syllable. The fact that the two tones go together and target the last two syllables of the word argues in favor of an analysis which treats this pattern as a property of the phrase edge, i.e., a boundary tone.

The votive chant melody can be seen in the following pitch tracks of trisyllabic words with the stress on the first syllable:¹⁰

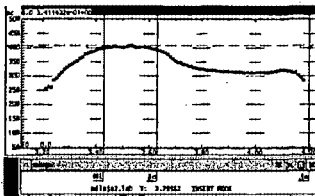


Figure 69: *Miloje!*

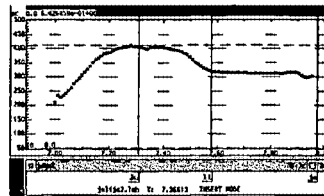


Figure 70: *Juliije!*

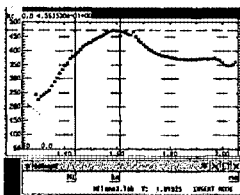


Figure 71: *Milane!*

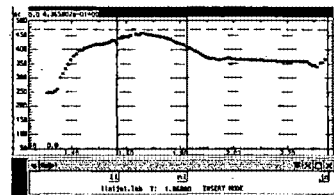


Figure 72: *linijo!*

¹⁰At this time I don't have an example of an all sonorant trisyllabic proper name under the long-rising accent, so I have used a common noun for illustration purposes.

It seems quite obvious that vocative chant is affecting the lexical specification for the tonal information. However, it is not entirely clear that the lexical information is completely lost. There seems to be at least one difference between falling and rising accents. Rising accents have a slightly higher target for the H tone in the above examples (compare Figures 69 and 70 with Figures 71 and 72). Whether the lexical H tone which correlates with the second syllable is boosting the boundary H tone is an open question and would require a detailed study.

To show that vocative chant is a boundary effect, we can look at examples of longer words. In the following two figures I show a calling contour on a word *Slobòdane!* and its possible extended variant, *Slobo-Slobòdane!*.

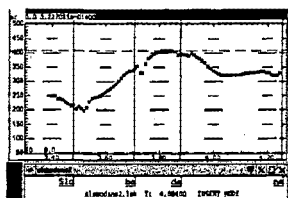


Figure 73: *Slobòdane!*

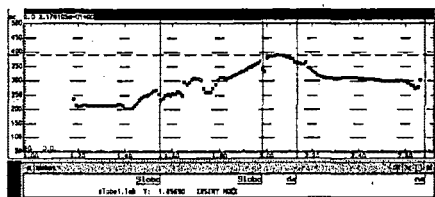


Figure 74: *Slobo-Slobòdane!*

F_0 peaks in the above contours occur on the penultimate syllable in both variants of summoning Slobodan. This shows that the vocative melody is truly a boundary effect.

To see that this bitonal boundary tone is targeting the last two syllables of a word regardless of the position of the stressed syllable, we can look at a stress initial word with more than three syllables. Consider the following pitch track.

SERBO-CROATIAN INTONATION

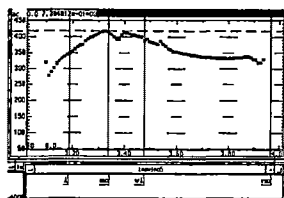


Figure 75: *imovino!*

Figure 75 shows again that the H tone is associated with the third, penultimate syllable, even though the stress is on the first syllable. Because the stress syllable does not play a role anchoring the melody, I conclude that vocative chant can be analyzed as a H- phrase accent followed by a L% boundary tone. This melody then differs from the prompting intonation where we saw that L- phrase accent is a special kind of boundary tone because it is timed to the stressed syllable.

There is one more thing to mention regarding vocative chant. The L tone of this boundary tone is not as low as the single L% boundary tone that we see in declarative utterances. It may seem reasonable then to question this characterization of this tone as a L. Presumably, another possible analysis for this contour would be to say that we have H- phrase accent followed by a downstepped !H% boundary tone. So far, we have no other evidence for a downstepped boundary tone. At this point then it seems unjustified to introduce a new target just for this melody. However, should such evidence arise, a reanalysis may be appropriate.

To sum up, in this section we have seen three different types of structural markers: L- phrase accent followed by a H% boundary tone (prompting intonation); a %H word boundary tone of double focus constructions, and a H- phrase accent followed by a L% boundary tone (vocative chant). It seems that the unifying property of these edge tones is that structural H tones affect lexical information so that lexical pitch accents end up neutralized. This is still a tentative conclusion until more data become available.

6 Conclusion

The surface tones of Serbo-Croatian accents can be described in terms of alternations between H and L tones, with some tones assigned in the lexicon, others assigned at the level of prosodic phrasing and yet others functioning to integrate pragmatic coherence of the discourse. I have argued that falling accents have a HL melody whereas rising accents have LH melody. The difference between the short and long falling is in the timing of the fall, which is the function of the length of the tone bearing unit (mora) to which the

H tone is anchored to. Long falling accents have the beginning of the realization of the fall on the stressed syllable itself and continuing on the following syllable. The fall of the short falling accent is delayed until the post-stressed syllable. However, for the purpose of the phonological representation it is sufficient to represent both falling accents as H^*+L , since the duration of the stressed syllable will determine the positioning of the trailing tone. Both rising accents show the LH melody where the L occurs on the stressed syllable and the H on the post-stressed syllable. Consequently, the phonological representation for the rising accents is L^*+H . The timing of the trailing H tone of the rising accents also does not need to be stipulated. Anchoring the L tone to the (last) mora of the accented syllable produces the desired effect of having the H tone on the post-stressed syllable and yet gives us enough flexibility, as with the falling accents, to accomodate variations in production.

The short/long distinction between the rising accents also seems to be accompanied by a difference in vowel quality. Thus it might be necessary to include a study of vowel quality together with the lexical accentual properties. Therefore, the proposal offered here for the description of the accents only in terms of a two way distinction, (falling vs. rising) may be necessary and sufficient.

A broad focus declarative utterance allows all lexical tones to be realized. Phonological words are clearly separated by %L, a word boundary tone. In addition, each subsequent phonological word is down-stepped from the previous one. The sentence initial constituent, regardless of its syntactic function, is set off from the rest of the constituents by having the highest target for the realization of the lexical H. The sentence final constituent is conversely in the lowest pitch range. Nevertheless, the realization of the lexical accents is still present. The falling accents in this position show a steady fall in the pitch, whereas the rising accents maintain the same pitch level in the post-stressed syllable, thus marking the two accents differently. Focusing the final constituent in a sentence allows all lexical tones to be fully realized, providing additional evidence that the phonological representation is not lost in final position.

In prosodic narrow focus utterances, the constituents following the focused constituent are in a markedly reduced pitch range relative to an utterance without the prosodic focus. In a paradigmatic contrast with broad focus utterances, prosodic focus slightly expands the pitch range of the focused constituent, creating a higher target for the lexical H, and compresses the pitch range surrounding the focused constituent, most drastically the following ones. That is, focal prominence involves not so much making the focal peak higher as it does make non-focal peaks lower. I have proposed that post-focal pitch range reduction is a consequence of the early realization of the L- phrase accent at the right edge of the focused word. The double focus construction provides evidence for a %H word boundary tone, which is used to signal the dependent variable constituent in this construction. This boundary tone differs from pitch range expansion in narrow focus constructions in the following way: pitch range expansions provides a wider tonal space for all tonal

SERBO-CROATIAN INTONATION

targets, %H word boundary tone raises the tonal target of the left edge of the word thereby creating a pull for the preceding and subsequent L tones.

Morphologically marked questions do not have a H% boundary tone, whereas non-marked questions can be signaled either by a question-tag which always has a H% boundary tone, or by prompting intonation on the focal constituent. Prompting intonation and vocative chant is a result of the combination of a phrase accent followed by a boundary tone, L- H% and H- L% respectively. Both of these intonational contours seem to affect lexical pitch accents.

To sum up: in this paper I have argued for three levels of prosodic phrasing in Serbo-Croatian, a phonological word, an intermediate phrase and an intonational phrase. The two prosodic units are either associated with certain tonal markings, such as edge tones, or function as a domain of a rule application. A phonological word has a delimitative marker, an initial wordy boundary tone, which can be either %L or %H, and a culminative marker, which can be any of the four pitch accents. The intermediate phrase is marked by phrasal accents (L- and H-) and the intonational phrase has two types of boundary tones (L% or H%).

Since Serbo-Croatian is both a stress language and a pitch accent language, it provides an example of a very different type of language than the ones that have been studied in depth so far from an intonational point of view, such as English, Japanese and others. In particular, I hope to have shown that one of the main points of interest in study of Serbo-Croatian intonation is the interaction of lexical tonal specifications with the tonal markings of intonational phrasing.

REFERENCES

- Beckman, M. and J. Pierrehumbert (1986). 'Intonational Structure in Japanese and English'. *Phonology Yearbook* 3:255-309.
- Brown, G., K. L. Currie, & J. Kennworthy (1980). *Questions of Intonation*. London: Croom Helm.
- Browne, E. W. (1967). 'On the Problem of Enclitic Placement in Serbo-Croatian'. MIT ms.
- Browne, E. W. and J. D. McCawley (1965). 'Srpskohrvatski akcenat'. *Zbornik za filologiju i lingvistiku* 8:147-151.
- Bruce, Gösta (1977). *Swedish word accents in sentence perspective*. Lund: Gleerup.
- Bruce, Gösta (1990). 'Alignment and composition of tonal accents: comments on Silverman and Pierrehumbert's paper', in Beckman M. and J. Kingston (1990) *Between the grammar and the physics of speech* Papers on Laboratory Phonology: 1. Cambridge: Cambridge University Press.
- Clements, G. N. and K. C. Ford (1979). 'Kikuyu Tone Shift', *LI* 10.2.
- Clements, G. N. and K. C. Ford (1981). 'On the Phonological Status of Downstep in Kikuyu', in D. L. Goyvaerts ed. (1981). *Phonology in the 1980's*. Ghent/Belgium: E. Story-Scientia.
- Godjevac, S. (1999). 'Declarative Utterance-Final Position in Serbo-Croatian' in Osamu Fujimura, Brian Joseph and Bogumil Palek (eds.) *Proceedings of LP'98*. Prague: The Karolinum Press. 63-76.
- Grice, M., R. D. Ladd, and A. Arvaniti (in press) "On the place of phrase accents in intonational phonology". *Phonology*. Cambridge: Cambridge University Press.
- Grønnum, N. (1985). 'Intonation and text in Standard Danish'. *JASA* 77(3):1205-1216.
- Grønnum, N. (1992). *The groundworks of Danish intonation: an introduction*. Copenhagen: University of Copenhagen/Museum Tusculanum Press.
- Gvozdanović (1980). *Tone and Accent in Standard Serbo-Croatian*. Verlag Der Österreichischen Akademie Der Wissenschaften Wien.
- Halle, M. (1971). 'Remarks on Slavic Accentology'. *LI* 1.2
- Harnsberger, J. D. and J. Judge (1996). 'Pitch Range and Focus in Hindi', paper given at 131st meeting of the Acoustical Society of America in Indianapolis.
- Hirschberg, J. & J. Pierrehumbert (1986). 'The Intonational Structuring of Discourse'. *Proceedings of the 24th Annual Meeting of the Association for Computational Linguistics*.

SERBO-CROATIAN INTONATION

- Inkelas, S. and D. Zec (1988). 'Serbo-Croatian Pitch Accent: The Interaction of Tone, Stress, and Intonation'. *Language* 64.2:227-248
- Inkelas, S. and D. Zec (1990), *The Phonology-Syntax Connection*, CSLI, The University of Chicago Press, Chicago.
- Jackendoff, R. (1972). *Semantic Interpretation in Generative Grammar*. Cambridge, Mass. MIT Press.
- Jin, Shunde (1996). *An Acoustic Study of Sentence Stress in Mandarin Chinese*. PhD dissertation, OSU.
- Jun, Sun-Ah & Mira Oh (1996). 'A Prosodic Analysis of Three Types of Wh-Phrases in Korean'. *Language and Speech* 39(1):37-61.
- Kariya, C. S. (1983). *The Acquisition of Accent in Serbo-Croatian: A Case Study*. PhD Thesis, Harvard University.
- Kostić, Dj. (1983). *Rečenička melodija u Srpskohrvatskom jeziku*. Beograd: Rad.
- Kubozono, H. (1992). 'Modeling syntactic effects on downstep in Japanese' in Docherty, G. J. & R. D. Ladd (eds.) *Gesture, segment, prosody - (Papers in laboratory phonology v.2)*. Cambridge: Cambridge University Press.
- Ladd, R. (1984). 'Declination: a review and some hypotheses' *Phonology Yearbook* 1:53-74.
- Ladd, R. (1988). 'Declination 'reset' and the hierarchical organization of utterances'. *JASA* 84:530-44.
- Ladd, R. (1996). *Intonational Phonology*. Cambridge: Cambridge University Press.
- Lehiste, I. (1975). 'The phonetic structure of paragraphs', in A. Cohen & S.G. Nooteboom (eds.) *Structure and Process in Speech Perception*. Berlin: Springer-Verlag, pp.195-206.
- Lehiste, I. and P. Ivić (1963). *Accent in Serbocroatian: An experimental Study*. Michigan Slavic Materials 4. Ann Arbor: University of Michigan, Dept. of Slavic Languages and Literatures.
- Lehiste, I. and P. Ivić (1977). 'Interrelationship between Word Tone and Sentence Intonation in Serbocroatian' in Donna Jo Napoli ed. *Elements of Tone, Stress, and Intonation*. Georgetown University Press, Washington, D.C.
- Lehiste, I. and P. Ivić (1986). *Word and Sentence Prosody in Serbocroatian*. Cambridge, Massachusetts: The MIT Press.
- Liberman, M. & J. Pierrehumbert (1984). 'Intonational Invariance under Changes in Pitch Range and Length' in M. Aronoff & R. T. Oerhle *Studies in Phonology. Presented to Morris Halle by his teacher and students* Cambridge, Massachusetts: The MIT Press.

- Maekawa, K. (1991). 'Perception of intonational characteristics of WH and non-WH questions in Japanese'. Proceedings of the XIIth International Congress of Phonetic Sciences, 4/5:202-205.
- Nespor, M. & I. Vogel. (1986). *Prosodic Phonology*. Dordrecht: Foris.
- Nikolić, B. (1970). *Osnovi Mladje Novoštokavske Akcentuacije*. Institut za Srpskohrvatski Jezik, Beograd.
- Pierrehumbert, J. (1980). *The phonetics and phonology of English intonation*. PhD dissertation, MIT.
- Pierrehumbert, J. & M. Beckman (1988). *Japanese Tone Structure*. Cambridge, Massachusetts: The MIT Press.
- Selkirk, E. (1984). *Phonology and Syntax: the Relation between Sound and Structure*. Cambridge, MA: MIT Press.
- Selkirk, E. (1986). "On derived domains in sentence phonology". *Phonology Yearbook* 3: 371-405.
- Stevanović, M. (1989). *Savremeni Srpskohrvatski Jezik*. Beograd: Naučna Knjiga.
- Venditti, J. (1994). 'The influence of syntax on prosodic structure in Japanese'. *OSU Working Papers in Linguistics* 44:191-223.